



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :		A1	(11) International Publication Number: <b>WO 97/44460</b> (43) International Publication Date: 27 November 1997 (27.11.97)
C12N 15/12, C07K 14/47, C12N 15/62, C07K 16/18, A61K 38/16, G01N 33/50, C12Q 1/68, C12N 1/21, 5/10, 5/12, A61K 48/00			
(21) International Application Number: PCT/US97/09303			(US). ICHIMURA, Takaharu [JP/US]; 32 Bradbury Street, Cambridge, MA 02138 (US). WEI, Henry [CN/US]; 62 Farwell Street, Newton, MA 02142 (US). CATE, Richard, L. [US/US]; 64 Arrowhead Road, Weston, MA 02193 (US).
(22) International Filing Date: 23 May 1997 (23.05.97)			
(30) Priority Data: 60/018,228 24 May 1996 (24.05.96) 60/023,442 23 August 1996 (23.08.96)		US US	(74) Agent: LEVINE, Leslie, M.; Biogen, Inc., 14 Cambridge Center, Cambridge, MA 02142 (US).
(60) Parent Application or Grant (63) Related by Continuation US 60/018,228 (CIP) Filed on 24 May 1996 (24.05.96)			(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(71) Applicant ( <i>for all designated States except US</i> ): BIOGEN, INC. [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).			
(72) Inventors; and (75) Inventors/Applicants ( <i>for US only</i> ): SANICOLA-NADEL, Michele [US/US]; 4 Maple Road, Winchester, MA 01890 (US). BONVENTRE, Joseph, V. [US/US]; 101 Boston Post Road, Wayland, MA 01778 (US). HESSION, Catherine, A. [US/US]; 35 Otis Hill Road, Hingham, MA 02043			Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: MODULATORS OF TISSUE REGENERATION			
(57) Abstract			
Proteins which are upregulated in injured or regenerating tissues, as well as the DNA encoding these proteins, are disclosed, as well as therapeutic compositions and methods of treatment encompassing these compounds.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## **MODULATORS OF TISSUE REGENERATION**

### **FIELD OF THE INVENTION**

The invention relates to proteins which are upregulated in injured or regenerating tissues, as well as to the DNA encoding these proteins. The invention further relates to therapeutic compositions and methods of treatment encompassing these proteins.

5

### **BACKGROUND OF THE INVENTION**

A dynamic remodeling of tissue architecture occurs during development and during tissue repair after injury. To study this process, we have focused on a model of kidney injury caused by an ischemia-reperfusion insult.

The kidney is able to repair damage to the proximal tubule epithelium through a complex series of events involving cell death, proliferation of surviving proximal tubule epithelial cells, formation of poorly differentiated regenerative epithelium over the denuded basement membrane, and differentiation of the regenerative epithelium to form a fully functional proximal tubule epithelial cells (Wallin et al., Lab. Invest. 66:474-484, 1992; Witzgall et al., Mol. Cell. Biol. 13:1933-1942, 1994; Ichimura et al., Am. J. Physiol. 269:F653-662, 1995; Thadhani et al., N. Engl. J. Med. 334:1448-1460, 1996). Growth factors such as IGF, EGF, and HGF have been implicated in this process of repair, as has the endothelial cell adhesion molecule ICAM-1. However, the mechanisms by which the tubular epithelial cells are restored are still not understood.

To identify molecules involved in process of injury and repair of the tubular epithelium, we analyzed the difference in the mRNA populations between injured/regenerating and normal kidneys using representational difference analysis (RDA). RDA is a PCR-based method for subtraction which yields target tissue or cell specific cDNA fragments by repetitive subtraction and amplification (Hubank and Schutz, Nucl. Acids Res. 22:5640-5648, 1994).

- 2 -

### SUMMARY OF THE INVENTION

The invention generally provides Kidney Injury-related Molecules (each of which is henceforth called a "KIM") which are upregulated in renal tissue after injury to the kidney. The KIM proteins and peptides of the invention, as well as their agonists and antagonists, and their corresponding are useful in a variety of therapeutic interventions.

The invention provides a purified and isolated DNA molecule having a nucleotide sequence set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6. The invention also includes the complementary strands of these sequences, DNA molecules which hybridize under stringent conditions to the aforementioned DNA molecules, and DNA molecules which, but for the degeneracy of the genetic code, would hybridize to any of the DNA molecules defined above. These DNA molecules may be recombinant, and may be operably linked to an expression control sequence.

The invention further provides a vector comprising a purified and isolated DNA molecule having a nucleotide sequence set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, or one of the other DNA molecules defined above. This vector may be a biologically functional plasmid or viral DNA vector. One embodiment of the invention provides a prokaryotic or eukaryotic host cell stably transformed or transfected by a vector comprising a DNA molecule of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6. In another embodiment of the invention, a process is provided for the production of a KIM polypeptide product encoded by a DNA molecule as described above; the process involves growing, under suitable culture conditions, prokaryotic or eukaryotic host cells transformed or transfected with the DNA molecule in a manner allowing expression of the DNA molecule, and recovering the polypeptide product of said expression.

A purified and isolated human KIM protein substantially free of other human proteins is specifically within the invention, as is a process for the production of a polypeptide product having part or all of the primary structural conformation and the biological activity of a KIM protein. KIM proteins of the invention may have an amino acid sequence which comprises SEQ ID NO:3, SEQ ID NO:5, or SEQ ID NO:7, or may be a variant of SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:7, or a purified and isolated protein encoded by the DNA of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6. These proteins can be provided substantially free of other human proteins. The invention further includes variants of these proteins, such as soluble

- 3 -

variants or fusion proteins. KIM fusion proteins of the invention may comprise an immunoglobulin, a toxin, an imageable compound or a radionuclide.

The invention also provides a specific monoclonal antibody to the KIM proteins described above. The anti-KIM antibody may be associated with a toxin, imageable compound or radionuclide. Further taught is a hybridoma cell line which produces such a specific antibody.

Pharmaceutical compositions are also within the scope of the invention. A pharmaceutical composition of the invention may comprise a therapeutically effective amount of a KIM protein or anti-KIM antibody of the invention, along with a pharmacologically acceptable carrier.

10 Diagnostic methods are within the invention, such as assessing the presence or course of resolution of renal injury by measuring the concentration of KIM in urine, serum, or urine sediment of patients who have or who are at risk of developing renal disease.

Methods of treatment of the invention include treating patients with therapeutically effective amounts of KIM, KIM variants, KIM analogs, KIM fusion proteins, KIM agonists, and 15 antibodies to KIM or to KIM ligands. Other therapeutic compounds of the invention include KIM ligands, anti-KIM antibodies, and fusions proteins of KIM ligands. These compounds can be useful in therapeutic methods which either stimulate or inhibit cellular responses that are dependent on KIM function.

Further methods of the invention inhibit the growth of KIM-expressing tumor cells by 20 contacting the cells with a fusion protein of a KIM ligand and either a toxin or radionuclide, or with an anti-KIM antibody conjugated to a toxin or to a radionuclide. Likewise, growth of tumor cells which express KIM ligand may be inhibited by contacting the cells with a fusion protein of a KIM and either a toxin or radionuclide, or with an anti-KIM ligand antibody conjugated to a toxin or to a radionuclide.

25 The invention also encompasses methods of gene therapy. These include a method of treating a subject with a renal disorder, a method of promoting growth of new tissue in a subject, and a method of promoting survival of damaged tissue in a subject, comprising administering to the subject a vector which includes DNA comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6.

30 The compounds of the invention are also useful for imaging tissues, either *in vitro* or *in vivo*. One such method involves targeting an imageable compound to a cell expressing a protein

- 4 -

of SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:7, comprising contacting the cell with either a monoclonal antibody of the invention or a fusion protein comprising a protein as described above, fused to an imageable compound. For *in vivo* methods, the cell is within a subject, and the protein or the monoclonal antibody is administered to the subject.

- 5       The invention also includes diagnostic methods, such as a method of identifying damage or regeneration of renal cells in a subject, comprising comparing the level of expression of either SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6 in renal cells of the subject to a control level of expression of the sequence in control renal cells. Another method of the invention includes identifying upregulation of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or  
10      SEQ ID NO:6 in cells comprising contacting the cells with an antisense probe and measuring hybridization to RNA within the cell.

A further embodiment of the diagnostic methods of the invention includes assessing the presence or concentration of a molecule of the invention either in urine, serum, or other body fluids, or in urine sediment or tissue samples. The measured injury-related molecule can be  
15      correlated with the presence, extent or course of a pathologic process. This correlation can also be used to assess the efficacy of a therapeutic regime.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 is the nucleotide sequence of rat clone cDNA 3-2, with putative protein reading frame of 615 to 1535.

- 20      FIGURE 2 is a listing of the cDNA sequence of rat clone 1-7, with putative protein reading frame of 145 to 1065.

FIGURE 3 is a listing of the cDNA sequence of rat clone 4-7, with putative protein reading frame of 107 to 1822.

- 25      FIGURE 4 is a listing of the cDNA and deduced amino acid sequences of human clone HI3-10-85, with putative protein reading frame of 1 to 1002. The upper line of the listing is the cDNA

- 5 -

sequence (SEQ ID NO:6), and the lower line is the deduced amino acid sequence (SEQ ID NO:7).

FIGURE 5 is a BESTFIT comparison of the nucleotide sequence of human clone H13-10-85 with that of rat clone 3-2.

5

#### DETAILED DESCRIPTION OF THE INVENTION

We identified KIM genes by analyzing differences in mRNA expression between regenerating and normal kidneys using representational difference analysis (RDA). RDA is a PCR-based method for subtraction which yields target tissue or cell-specific cDNA fragments by repetitive subtraction and amplification. The cDNA representation from 48 hr postischemic 10 adult rat kidney RNA is subtracted with the sample from normal (sham-operated) adult rat kidney. In this procedure, sequences which are common to both postischemic and to normal kidney samples are removed, leaving those sequences which are significantly expressed only in the injured kidney tissue. Such genes encode proteins that may be therapeutically beneficial for renal disorders or involved in the injury process. Several clones have been obtained, sequenced 15 and characterized. The clones are then investigated for their expression patterns during kidney repair, development and tissue distribution by northern analysis and RNA *in situ* hybridization.

#### Sequence Identification Numbers

Nucleotide and amino acid sequences referred to in the specification have been given the following sequence identification numbers:

- 20 SEQ ID NO:1 - nucleotide sequence of rat 3-2 cDNA insert  
SEQ ID NO:2 - nucleotide sequence of rat 1-7 cDNA insert  
SEQ ID NO:3 - amino acid sequence of rat KIM-1, encoded by rat 3-2 and 1-7 cDNA's  
SEQ ID NO:4 - nucleotide sequence of rat 4-7 cDNA insert  
SEQ ID NO:5 - amino acid sequence encoded by 4-7 cDNA insert  
25 SEQ ID NO:6 - nucleotide sequence of human cDNA clone H13-10-85  
SEQ ID NO:7 - amino acid sequence encoded by human cDNA clone H13-10-85

- 6 -

**Definitions of Terms**

A "KIM protein", herein used synonymously with "KIM", is a protein encoded by mRNA which is selectively upregulated following injury to a kidney. One group of KIM proteins of interest includes those coded for by mRNA which is selectively upregulated at any time within one week following any insult which results in injury to renal tissue. Examples of times at which such upregulation might be identified include 10 hours, 24 hours, 48 hours or 96 hours following an insult. Examples of types of insults include those resulting in ischemic, toxic or other types of injury.

A "KIM agonist" is a molecule which can specifically trigger a cellular response normally triggered by the interaction of KIM with a KIM ligand. A KIM agonist can be a KIM variant, or a specific antibody to KIM, or a soluble form of the KIM ligand.

A "KIM antagonist" is a molecule which can specifically associate with a KIM ligand or KIM, thereby blocking or otherwise inhibiting KIM binding to the KIM ligand. The antagonist binding blocks or inhibits cellular responses which would otherwise be triggered by ligation of the KIM ligand with KIM or with a KIM agonist. Examples of KIM antagonists include certain KIM variants, KIM fusion proteins and specific antibodies to a KIM ligand or KIM.

A "KIM ligand" is any molecule which noncovalently and specifically binds to a KIM protein. Such a ligand can be a protein, peptide, steroid, antibody, amino acid derivative, or other type molecule, in any form, including naturally-occurring, recombinantly produced, or otherwise synthetic. A KIM ligand can be in any form, including soluble, membrane-bound, or part of a fusion construct with immunoglobulin, fatty acid, or other moieties. The KIM ligand may be an integrin. A membrane-bound KIM ligand can act as a receptor which, when bound to or associated with KIM, triggers a cellular response. In some interactions, KIM may associate with more than a single KIM ligand, or may associate with a KIM ligand as part of a complex with one or more other molecules or cofactors. In a situation where both the KIM and the KIM ligand are bound to cell membranes, the KIM may associate and react with KIM ligand which is bound to the same cell as the KIM, or it may associate and react with KIM ligand be bound to a second cell. Where the KIM ligation occurs between molecules bound to different cells, the two cells may be the same or different with respect to cellular type or origin, phenotypic or metabolic condition, or type or degree of cellular response (e.g., growth, differentiation or apoptosis) to a given stimulus. "KIM ligation" refers to the contact and binding of KIM with a KIM ligand.

- By "alignment of sequences" is meant the positioning of one sequence, either nucleotide or amino acid, with that of another, to allow a comparison of the sequence of relevant portions of one with that of the other. An example of one method of this procedure is given in Needleman et al. (J. Mol. Biol. 48:443-453, 1970). The method may be implemented conveniently by
- 5 computer programs such as the Align program (DNAstar, Inc.). As will be understood by those skilled in the art, homologous or functionally equivalent sequences include functionally equivalent arrangements of the cysteine residues within the conserved cysteine skeleton, including amino acid insertions or deletions which alter the linear arrangement of these cysteines, but do not materially impair their relationship in the folded structure of the protein. Therefore,
- 10 internal gaps and amino acid insertions in the candidate sequence are ignored for purposes of calculating the level of amino acid sequence homology or identity between the candidate and reference sequences. One characteristic frequently used in establishing the homology of proteins is the similarity of the number and location of the cysteine residues between one protein and another.
- 15 "Antisense DNA" refers to the sequence of chromosomal DNA that is transcribed. An "antisense probe" is a probe which comprises at least a portion of the antisense DNA for a nucleic acid portion of interest.
- By "cloning" is meant the use of *in vitro* recombination techniques to insert a particular gene or other DNA sequence into a vector molecule. In order to successfully clone a desired
- 20 gene, it is necessary to employ methods for generating DNA fragments, for joining the fragments to vector molecules, for introducing the composite DNA molecule into a host cell in which it can replicate, and for selecting the clone having the target gene from amongst the recipient host cells.
- By "cDNA" is meant complementary or copy DNA produced from an RNA template by the action of RNA-dependent DNA polymerase (reverse transcriptase). Thus a "cDNA clone"
- 25 means a duplex DNA sequence complementary to an RNA molecule of interest, carried in a cloning vector.
- By "cDNA library" is meant a collection of recombinant DNA molecules containing cDNA inserts which together comprise a representation of the mRNA molecules present in an entire organism or tissue, depending on the source of the RNA templates. Such a cDNA library
- 30 may be prepared by methods known to those of skill, and described, for example, in Maniatis et al., Molecular Cloning: A Laboratory Manual, *supra*. Generally, RNA is first isolated from the

- 8 -

cells of an organism from whose genome it is desired to clone a particular gene. Preferred for the purposes of the present invention are mammalian, and particularly human, cell lines. Alternatively, RNA may be isolated from a tumor cell, derived from an animal tumor, and preferably from a human tumor. Thus, a library may be prepared from, for example, a human 5 adrenal tumor, but any tumor may be used.

As used herein, the term "DNA polymorphism" refers to the condition in which two or more different nucleotide sequences can exist at a particular site in DNA.

"Expression vector" includes vectors which are capable of expressing DNA sequences contained therein, i.e., the coding sequences are operably linked to other sequences capable of 10 effecting their expression. It is implied, although not always explicitly stated, that these expression vectors must be replicable in the host organisms either as episomes or as an integral part of the chromosomal DNA. A useful, but not a necessary, element of an effective expression vector is a marker encoding sequence, which is a sequence encoding a protein which results in a phenotypic property (such as tetracycline resistance) of the cells containing the protein which 15 permits those cells to be readily identified. In sum, "expression vector" is given a functional definition, and any DNA sequence which is capable of effecting expression of a specified contained DNA code is included in this term, as it is applied to the specified sequence. As at present, such vectors are frequently in the form of plasmids, thus "plasmid" and "expression vector" are often used interchangeably. However, the invention is intended to include such other 20 forms of expression vectors which serve equivalent functions and which may, from time to time become known in the art.

By "functional derivative" is meant the "fragments", "variants", "analogs", or "chemical derivatives" of a molecule. A "fragment" of a molecule, such as any of the antigens of the present invention is meant to refer to any polypeptide subset of the molecule. A "variant" of 25 such molecules is meant to refer to a naturally occurring molecule substantially similar to either the entire molecule, or a fragment thereof. An "analog" of a molecule is meant to refer to a non-natural molecule substantially similar to either the entire molecule or a fragment thereof.

The term "gene" means a polynucleotide sequence encoding a peptide.

By "homogeneous" is meant, when referring to a peptide or DNA sequence, that the 30 primary molecular structure (i.e., the sequence of amino acids or nucleotides) of substantially all molecules present in the composition under consideration is identical.

- 9 -

"Isolated" refers to a protein of the present invention, or any gene encoding any such protein, which is essentially free of other proteins or genes, respectively, or of other contaminants with which it might normally be found in nature, and as such exists in a form not found in nature.

5        The term "label" refers to a molecular moiety capable of detection including, by way of example, without limitation, radioactive isotopes, enzymes, luminescent agents, and dyes.

The term "probe" refers to a ligand of known qualities capable of selectively binding to a target antiligand. As applied to nucleic acids, the term "probe" refers to a strand of nucleic acid having a base sequence complementary to a target strand.

10      "Recombinant host cells" refers to cells which have been transformed with vectors constructed using recombinant DNA techniques. As defined herein, the antibody or modification thereof produced by a recombinant host cell is by virtue of this transformation, rather than in such lesser amounts, or more commonly, in such less than detectable amounts, as would be produced by the untransformed host.

15      By "substantially pure" is meant any protein of the present invention, or any gene encoding any such protein, which is essentially free of other proteins or genes, respectively, or of other contaminants with which it might normally be found in nature, and as such exists in a form not found in nature.

A molecule is said to be "substantially similar" to another molecule if the sequence of 20 amino acids in both molecules is substantially the same, and if both molecules possess a similar biological activity. Thus, provided that two molecules possess a similar activity, they are considered variants as that term is used herein even if one of the molecules contains additional amino acid residues not found in the other, or if the sequence of amino acid residues is not identical. As used herein, a molecule is said to be a "chemical derivative" of another molecule 25 when it contains additional chemical moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 16th ed., Mack Publishing Co.,  
30 Easton, Penn. (1980).

- 10 -

By "vector" is meant a DNA molecule, derived from a plasmid or bacteriophage, into which fragments of DNA may be inserted or cloned. A vector will contain one or more unique restriction sites, and may be capable of autonomous replication in a defined host or vehicle organism such that the cloned sequence is reproducible.

5    **Compounds of the Invention**

The invention includes the cDNA of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6, as well as sequences which include the sequence of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6, and derivatives of these sequences. The invention also includes vectors, liposomes and other carrier vehicles which encompass these sequence or 10 derivatives of them. The invention further includes proteins transcribed from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6, including but not limited to SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:7, and their derivatives and variants.

One embodiment of the invention includes soluble variants of a KIM protein that is usually synthesized as a membrane associated protein, and which is upregulated after injury. 15 Soluble variants lack at least a portion of the transmembrane or intra-membrane section of a native KIM protein. In some examples, the soluble variant lacks the entire transmembrane or intra-membrane section of a native KIM protein. Soluble variants include fusion proteins which encompass derivatives of KIM proteins that lack at least a portion of the transmembrane or intra-membrane section of a native KIM protein. All types of KIM fusion proteins are included, 20 particularly those which incorporate his-tag, Ig-tag, and myc-tag forms of the molecule. These KIM fusions may have characteristics which are therapeutically advantageous, such as the increased half-life conferred by the Ig-tag. Also included are fusion proteins which incorporate portions of selected domains of the KIM protein.

Variants can differ from naturally occurring KIM protein in amino acid sequence or in 25 ways that do not involve sequence, or both. Variants in amino acid sequence are produced when one or more amino acids in naturally occurring KIM protein is substituted with a different natural amino acid, an amino acid derivative or non-native amino acid. Particularly preferred variants include naturally occurring KIM protein, or biologically active fragments of naturally occurring KIM protein, whose sequences differ from the wild type sequence by one or more conservative 30 amino acid substitutions, which typically have minimal influence on the secondary structure and

- 11 -

hydrophobic nature of the protein or peptide. Variants may also have sequences which differ by one or more non-conservative amino acid substitutions, deletions or insertions which do not abolish the KIM protein biological activity. Conservative substitutions typically include the substitution of one amino acid for another with similar characteristics such as substitutions 5 within the following groups: valine, glycine; glycine, alanine; valine, isoleucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. The non-polar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine and glutamine. The positively charged 10 (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Other conservative substitutions can be taken from the table below, and yet others are described by Dayhoff in the Atlas of Protein Sequence and Structure (1988).

- 12 -

**TABLE 1: CONSERVATIVE AMINO ACID REPLACEMENTS**

For Amino Acid	Code	Replace with any of
Alanine	A	D-Ala, Gly,beta-Ala, L-Cys,D-Cys
Arginine	R	D-Arg, Lys,homo-Arg, D-homo-Arg, Met,D-Met, Ile, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn,Asp,D-Asp,Glu,D-Glu, Gln,D-Gln
Aspartic Acid	D	D-Asp,D-Asn,Asn, Glu,D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys,Met,D-Met,Thr, D-Thr
Glutamine	Q	D-Gln,Asn, D-Asn,Glu,D-Glu,Asp, D-Asp
Glutamic Acid	E	D-Glu,D-Asp,Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala,Pro, D-Pro, Beta-Ala, Acp
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Met, D-Met
Lysine	K	D-Lys,Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D_Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val, Norleu
Phenylalanine	F	D-Phe,Tyr, D-Thr,L-Dopa,His,D-His, Trp, D-Trp, Trans 3,4 or 5-phenylproline, cis 3,4 or 5 phenylproline

- 13 -

Proline	P	D-Pro, L-I-thioazolidine-4-carboxylic acid, D- or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
5      Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

- 14 -

Other variants within the invention are those with modifications which increase peptide stability. Such variants may contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the peptide sequence. Also included are: variants that include residues other than naturally occurring L-amino acids, such as D-amino acids or non-naturally occurring 5 or synthetic amino acids such as beta or gamma amino acids and cyclic variants. Incorporation of D- instead of L-amino acids into the polypeptide may increase its resistance to proteases. See, e.g., U.S. Patent 5,219,990.

Generally, substitutions that may be expected to induce changes in the functional properties of KIM polypeptides are those in which: (i) a hydrophilic residue, e.g., serine or 10 threonine, is substituted by a hydrophobic residue, e.g., leucine, isoleucine, phenylalanine, or alanine; (ii) a cysteine residue is substituted for (or by) any other residue; (iii) a residue having an electropositive side chain, e.g., lysine, arginine or histidine, is substituted for (or by) a residue having an electronegative charge, e.g., glutamic acid or aspartic acid; or (iv) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having such a side chain, 15 e.g., glycine.

The peptides of this invention may also be modified by various changes such as insertions, deletions and substitutions, either conservative or nonconservative where such changes might provide for certain advantages in their use. Splice variants are specifically included in the invention.

In other embodiments, variants with amino acid substitutions which are less conservative 20 may also result in desired derivatives, e.g., by causing changes in charge, conformation and other biological properties. Such substitutions would include for example, substitution of hydrophilic residue for a hydrophobic residue, substitution of a cysteine or proline for another residue, substitution of a residue having a small side chain for a residue having a bulky side chain or 25 substitution of a residue having a net positive charge for a residue having a net negative charge.

When the result of a given substitution cannot be predicted with certainty, the derivatives may be readily assayed according to the methods disclosed herein to determine the presence or absence of the desired characteristics.

Variants within the scope of the invention include proteins and peptides with amino acid 30 sequences having at least eighty percent homology with a KIM protein. More preferably the sequence homology is at least ninety percent, or at least ninety-five percent. For the purposes

- 15 -

of determining homology the length of comparison sequences will generally be at least 8 amino acid residues, usually at least 20 amino acid residues. Variants of the compounds of the invention also includes any protein which 1) has an amino acid sequence which is at least forty percent homologous to a KIM protein of the invention, and also which 2) after being placed in an optimal alignment with the KIM sequence (as depicted in Figure 5 for human and for rat KIM-1) has at least 80% of its cysteine residues aligned with cysteines in the KIM protein of the invention.

- Just as it is possible to replace substituents of the scaffold, it is also possible to substitute functional groups which are bound to the scaffold with groups characterized by similar features.
- 10 These substitutions will initially be conservative, i.e., the replacement group will have approximately the same size, shape, hydrophobicity and charge as the original group. Non-sequence modifications may include, for example, *in vivo* or *in vitro* chemical derivatization of portions of naturally occurring KIM protein, as well as changes in acetylation, methylation, phosphorylation, carboxylation or glycosylation.
- 15 Also included within the invention are agents which specifically bind to the protein, or a fragment of the protein (SEQ ID NO:3, 5 or 7). These agents include ligands and antibodies (including monoclonal, single chain, double chain, Fab fragments, and others, whether native, human, humanized, primatized, or chimeric). Additional descriptions of these categories of agents are in PCT application 95/16709, the specification of which is herein incorporated by reference.
- 20

### Experimental Procedures

#### 1. Generation of RNA from ischemic and normal rat adult kidneys

Ischemic injured rat kidneys are generated as described by Witzgall et al. (J. Clin Invest. 25: 93: 2175-2188, 1994). Briefly, the renal artery and vein from one kidney of an adult Sprague-Dawley rat are clamped for 40 minutes and then reperfused. Injured kidneys are harvested from the rats at 24 hours and at 48 hours after reperfusion. Kidneys from sham-operated, normal adult Sprague-Dawley rats are also harvested.

Total RNA is prepared from the organs based on the protocol by Glisin et al. (Biochemistry 13: 2633, 1974). Briefly, the harvested organs are placed immediately into GNC

- 16 -

buffer (4M guanidine thiocyanate, 0.5% SDS, 25mM sodium citrate, 0.1% Sigma anti foam) and disrupted on ice with a polytron. Cell debris is removed with a low speed spin in a clinical centrifuge and the supernatant fluid is placed on a 5.7 M CsCl, 25mM sodium acetate, 1mM EDTA cushion. RNA is pelleted through the cushion in a SW40Ti rotor at 22K for 15hrs. RNA 5 is resuspended in sterile DEPC- treated water, precipitated twice with 1/10 volume 3M sodium acetate and 2.5 volumes of EtOH. Poly A+ RNA is isolated using an mRNA purification kit (Pharmacia, catalog No.27-9258-02).

**2. Representational Difference Analysis (RDA) method to isolate 1-7, 3-2 and 4-7 RDA fragments**

10 Double stranded cDNA is synthesized from sham-operated and from 48hr post-ischemic kidney poly A+ RNA using Gibco BRL "Superscript Choice™ System cDNA Synthesis Kit", catalog No. 18090. First strand is synthesized by priming with oligo dT and using Superscript II™ reverse transcriptase. Second strand is generated using E. coli DNA polymerase I and RNase H followed by T4 DNA polymerase using BRL recommended conditions.

15 RDA analysis is performed essentially as described by Hubank and Schatz (Nucleic Acid Research 22: 5640-48, 1994). Briefly, 48 hr post-ischemic kidney cDNA is digested with the restriction enzyme *Dpn* II, and ligated to R-Bgl-12/ 24 oligonucleotides (see reference for exact sequence). PCR amplification (performed with Perkin-Elmer Taq polymerase and their corresponding PCR buffer) of the linker ligated cDNA is used to generate the initial 20 representation. This PCR product is designated "tester amplicon." The same procedure is used to generate "driver amplicon" from sham-operated rat kidney cDNA.

Hybridization of tester and driver amplicons followed by selective amplification are performed three times to generate Differential Product One (DP1), Two (DP2) and Three (DP3). Generation of the DP1 product is performed as described by Hubank and Schatz (Nucleic Acid 25 Research 22: 5640-48, 1994). The DP2 and DP3 products are also generated as described by Hubank and Schatz (id.), except that the driver:tester ratios are changed to 5,333:1 for DP2 and to 40,000 :1 or 4,000:1 for DP3.

Three RDA products are cloned from DP3 into the cloning vector pUC 18: RDA product 1-7 (252bp) when the DP3 was generated using a ratio of 40,000:1, and product RDA 3-2 30 (445bp) and 4-7 (483bp) when the DP3 was generated using a ratio of 4,000:1. The DNA

- 17 -

fragments are subcloned using the Pharmacia Sureclone™ kit (catalog No. 27-9300-01) to repair the ends of the PCR fragments with Klenow enzyme and to facilitate blunt end ligation of the fragments into the pUC18 vector.

### 3. Northern Analysis

5 Poly A+ RNA (2.5µg) from rat normal adult kidney (sham operated), from 48hr post-ischemic injured adult kidney, and from day 18 embryonic kidney is electrophoresed and Northern blotted (Cate, Cell 45:685, 1986) to a GeneScreen™ membrane (Dupont). Hybridization in PSB buffer (50mM Tris 7.5, 1M NaCl, 0.1% Na pyrophosphate, 0.2% PVP, 0.2% Ficoll, 0.2% BSA, 1% SDS), containing 10% dextran sulphate and 100µg/ml tRNA, is  
10 performed at 65°C using three different probes: 1-7 RDA product, 3-2 RDA product and 4-7 RDA product. All are radiolabeled using Pharmacia's "Ready to Go™" random priming labeling kit (catalog No.27-9251-01). RDA products 1-7, 3-2 and 4-7 hybridize to mRNAs present in all three samples, but most intensely to mRNAs in the 48hr post-ischemic kidney RNA samples.

A Northern blot analysis of adult rat tissues indicates that the 1-7 gene is expressed at  
15 very low levels in normal adult kidney, testis, spleen and lung. The 3-2 gene is expressed in liver, kidney, spleen, and brain. The 4-7 gene is expressed in spleen, kidney, lung, testis, heart, brain, liver, and skeletal muscle. The presence of different sized mRNAs in some tissues in the 1-7 and 3-2 blot indicates that the primary transcription product of the 1-7 gene and of the 3-2 gene may undergo alternate splicing and/or polyadenylation.

### 20 4. Isolation of 3-2 and 4-7 cDNA clones

A cDNA library is generated from 4 µg of polyA+ RNA from 48hr post-ischemic injured kidney using reagents from BRL Superscript Choice™ System for cDNA synthesis, and Stratagene™ Lambda ZapII cloning kit (catalog No. 236201), according to protocols recommended by the manufacturers.

25  $10^5$  clones are screened with the 3-2 RDA product as a probe (random primed labeled as described above). Eight positive clones are selected and four are randomly chosen for secondary analysis to obtain pure phage plaques. After tertiary screening, four pure phage clones are isolated. Cloned inserts from the phage are isolated by *in vivo* excision procedure according to Stratagene™ Lambda Zap II kit. The largest insert, of approximately 2.6 kb (referred to as

- 18 -

cDNA clone 3-2), is subjected to DNA sequencing. The sequence of the insert (SEQ ID NO:1) is shown in Figure 1. cDNA clone 3-2 (*E. coli* K-12, SOLR/p3-2#5-1) has been deposited as ATCC No. 98061. The sequence of cDNA clone 3-2 is identical to that of clone 1-7 cDNA (SEQ ID NO: 2), except that nucleotides 136-605 of SEQ ID NO:1 represent an insertion. Thus, 5 SEQ ID NO:2 represents a splice variant form of SEQ ID NO: 1. The clone for 1-7 (*E. coli* K-12, SOLR/p1-7#3-1) has been deposited as ATCC No. 98060.

10<sup>5</sup> clones are screened with the 1-7 RDA product as a probe (random primed radiolabeled as described above). Eight positive clones are selected and four are randomly chosen for secondary analysis to obtain pure phage plaques. After tertiary screening, four pure 10 phage clones are isolated. Cloned inserts from the phage are isolated by *in vivo* excision procedure according to Stratagene™ Lambda Zap II kit. The largest insert of approximately 2.0 kb (referred to as cDNA clone 1-7) is subjected to DNA sequencing; the sequence of the insert (SEQ ID NO: 2) is shown in Figure 2.

15 10<sup>5</sup> clones are screened with the 4-7 RDA product as a probe (random primed labeled as described above and hybridized in PSB at 65C). Eight positive clones are selected and four are randomly chosen for secondary analysis to obtain pure phage plaques. After secondary screening, two pure phage clones are isolated. Cloned inserts from the phage are isolated by *in vivo* excision procedure according to Stratagene™ Lambda Zap II kit. The largest insert, approximately 2.4 kb (referred to as cDNA clone 4-7), is subjected to DNA sequencing. The 20 sequence of the insert, SEQ ID NO: 4, is shown in Figure 3. The cDNA clone 4-7 (*E. coli* K-12, SOLR/p4-7#1-1) has been deposited as ATCC No. 98062..

##### 5. Characterization of the 1-7, 3-2 and 4-7 cDNA clones

###### A.) DNA and Protein Sequences:

25 The sequence of 3-2 cDNA (Figure 1; SEQ ID NO:1) contains an open reading frame of 307 amino acids (Figure 1; SEQ ID NO:3). A signal sequence of 21 amino acids is inferred from Von Heijne analysis (Von Heijne et al., Nucl. Acid Res. 14:14683 (1986)), and a transmembrane region spanning approximately aa 235-257 indicates that the 3-2 product is a cell surface protein.

30 The sequence of 1-7 cDNA (Figure 2; SEQ ID NO:2) contains an open reading frame of 307 amino acids, which is identical to the open reading frame contained in the 3-2 cDNA (SEQ ID NO: 3). The sequence of 4-7 cDNA (Figure 3; SEQ ID NO:4) contains an open reading

- 19 -

frame of 572 amino acids (SEQ ID NO:5). A transmembrane region is located at approximately amino acids 501-521.

B.) *In situ* analysis of 1-7, 3-2 and 4-7 mRNAs in contralateral and in post-ischemic adult rat kidneys:

- 5        *In situ* hybridization is carried out according to the method described by Finch et al., Dev. Dynamics 203: 223-240, 1995. Briefly, both ischemic and contralateral kidneys are perfusion fixed with 4% paraformaldehyde in PBS. Kidneys are further fixed overnight at 4C and processed. Paraffin sections are deparaffinized and rehydrated, fixed with 4% paraformaldehyde in PBS, digested with proteinase K, refixed, then acetylated with acetic anhydride in
- 10      triethanolamine buffer. Sections are then dehydrated and hybridized with <sup>32</sup>P-labeled riboprobes at 55 °C overnight, with <sup>33</sup>P-labeled riboprobes generated from 3-2 RDA or 1-7 RDA products subcloned into BamH1 site of pGEM-11Z. After hybridization, sections were washed under high stringency conditions (2 X SSC, 50 % formamide at 65°C). Sections are finally dehydrated, emulsion (NBT-2) coated for autoradiography, and exposed for at least a week. Silver grains are
- 15      developed and sections are counterstained with toluidine blue and microphotographed.

Analysis of 1-7 and 3-2 mRNA expression by *in situ* hybridization indicates that these genes are greatly upregulated in damaged kidney cells compared to their expression in normal kidney sections. The expression seen is in regenerative cells of the cortex and outer medulla, most of which appear to be proximal tubule cells.

- 20       Analysis of the 4-7 *in situ* RNA expression pattern also reveals abundant expression of this gene in the injured ischemic kidney compared to the normal adult kidney. The site of expression appears to be infiltrating cells.

6.) Isolation of a human cDNA clone which cross hybridizes to the rat 3-2 cDNA

- A <sup>32</sup>P-labeled DNA probe comprising nucleotides 546-969 of the insert of clone 3-2
- 25      shown in Figure 1 is generated and used to screen a human embryonic liver lambda gt10 cDNA library (Clontech Catalog #HL5003a). 1 X10<sup>6</sup> plaques are screened in duplicate using standard conditions as described above but temperature for screening was 55C. For the high stringency wash, the filters are washed in 2X SSC at 55C. Fifty positive phage are identified and plaque

- 20 -

purified, and DNA is prepared. The phage DNAs are subjected to Southern analysis using the same probe as above. The Southern blot filter is subjected to a final wash with 0.5X SSC at 55C. Two clones are identified as positive. The insert of clone H13-10-85 is sequenced and a region is found that encodes a protein with a high level of identity to the 3-2 protein shown in Figure 3.

5       The nucleotide sequence (SEQ ID NO:6) and predicted amino acid sequence (SEQ ID NO:7) of the human 3-2 related protein are shown in Figure 4. As shown by the bestfit analysis depicted in Figure 5, the human 3-2 related protein is 43.8% identical and 59.1% similar to the rat 3-2 protein. Both contain IgG, mucin, transmembrane, and cytoplasmic domains. The six cysteines within the IgG domains of both proteins are conserved.

10      7) Production of KIM-1 Ig fusion protein

A fusion protein of the extracellular domain of KIM and the Fc region of immunoglobulin (Ig) is a useful tool for the study of the molecular and cellular biology of the injured/regenerating kidney and as a therapeutic molecule. To produce Kim Ig fusion protein with the extracellular domain of human and rat KIM-1 protein, a fragment of the extracellular 15 domain of KIM-1 cDNA was amplified by PCR and cloned in the Biogen expression vector, pCA125, for transient expression in COS cells. The expression vector pCA125 produces a fusion protein which has a structure from gene cloned at N-terminus and a human Ig Fc region at the C-terminus. COS cells were transfected with the plasmids SJR 103 or 104; these plasmids express a fusion protein which contains the human KIM sequences 263-1147 (SEQ ID NO:6; SJR 103) or 20 rat KIM sequences 599-1319 (SEQ ID NO:1; SJR 104) of the extracellular domain fused to human Ig Fc region. The cells were grown in 10% FBS in DMEM in the cell factory (Nunc, Naperville, IL). Two to three days post-transfection, medium was harvested, concentrated using Amicon concentrator, and fusion protein was purified using Protein-A Sepharose column. After purification, purity of fusion protein was evaluated by SDS-PAGE.

25      Diagnostic Uses of the Compounds of the Invention

Anti-KIM antibodies of the invention, which specifically bind to the protein of SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:7 or a fragment thereof, are useful in several diagnostic methods. These agents may be labeled with detectable markers, such as fluoroscopically or radiographically opaque substances, and administered to a subject to allow imaging of tissues

which express KIM protein. The agents may also be bound to substances, such as horseradish peroxidase, which can be used as immunocytochemical stains to allow visualization of areas of KIM protein-positive cells on histological sections. A specific antibody could be used alone in this manner, and sites where it is bound can be visualized in a sandwich assay using an anti-  
5 immunoglobulin antibody which is itself bound to a detectable marker.

Specific antibodies to KIM protein are also useful in immunoassays to measure KIM presence or concentration in samples of body tissues and fluids. Such concentrations may be correlated with different disease states. As an embodiment of particular interest, the invention includes a method of diagnosing renal injury, or of monitoring a process of renal repair, by  
10 measuring the concentration of KIM or of KIM fragments in the urine, plasma or serum of a patient. Similarly, KIM can be measured in urine sediment, in particular in cellular debris in the urine sediment. Casts of renal tubule cells, which may be present in urine sediment from patients with ongoing renal disease, may contain elevated levels of KIM protein and mRNA.

Specific antibodies to KIM protein may also be bound to solid supports, such as beads or  
15 dishes, and used to remove the ligand from a solution, either for measurement, or for purification and characterization of the protein or its attributes (such as posttranslational modifications). Such characterization of a patient's KIM protein might be useful in identifying deleterious mutants or processing defects which interfere with KIM function and are associated with abnormal patient phenotypes. Each of these techniques is routine to those of skill in the  
20 immunological arts.

Additional imaging methods utilize KIM or KIM fragments, fused to imageable moieties, for diagnostic imaging of tissues that express KIM ligands, particularly tumors.

Further diagnostic techniques are based on demonstration of upregulated KIM mRNA in tissues, as an indication of injury-related processes. This technique has been tested and found  
25 workable in a model of ischemic injury in rats, as follows.

To determine if the amount of KIM-1 protein is increased after injury, we examined kidney homogenates of contralateral and postischemic kidneys 24 and 48 hours following a 40 minute clamping of the renal artery and vein of a single kidney for each rat. The kidney homogenate was assessed for the presence of KIM-1 protein. Western blot analysis identifies  
30 three proteins detected by two different antibodies after ischemic injury, which are not detectable in homogenates from contralateral kidneys which were not exposed to ischemic injury. The

- 22 -

apparent molecular weights of the bands are approximately 40kDa, 50kDa and 70-80kDa. The three protein species detected by western blotting could represent glycosylated forms of the same protein given the presence of potential N and O linked glycosylation sites. The fact that each of these proteins react with two different sets of polyclonal antibodies supports the idea that they  
5 are related to KIM-1 and are not cross-reacting bands. Confirmation of this prediction came from the results of partial CNBr cleavage of the three proteins which revealed they shared common CNBr cleavage fragments. Since the cytoplasmic domain of the KIM-1 protein is not predicted to contain any major post-translational modifications, the two smallest products of the digest (4.7kDa and 7.4kDa) detected with antibodies directed against the cytoplasmic domain of  
10 KIM-1 should be the same size for the three different KIM-1 protein bands if they originate from the same protein. We observed that the KIM1 40kDa and 70-80kDa proteins yield fragments migrating at the predicted size. Digest of the 50kDa protein band gave also the same C-terminal signature band peptide.

The KIM-1 sequence presents two putative sites for N-glycosylation and a mucin domain  
15 where O-glycosylation could cover the polypeptide chain. The three KIM-1 bands detected in postischemic kidney could correspond to glycosylation variants of the same core protein. De-N-glycosylation with PNGase F resulted in a shift of all three bands to a lower molecular weight, corresponding to a loss of about 3kDa, indicating that all three proteins are N-glycosylated. Differences in O-glycosylation might explain the differences in sizes of these three bands.

20 **Therapeutic Uses of the Compounds of the Invention**

The therapeutic methods of the invention involve selectively promoting or inhibiting cellular responses that are dependent on KIM ligation. Where the KIM and the KIM ligand are both membrane bound, and expressed by different cells, the signal transduction may occur in the KIM-expressing cell, in the KIM ligand-expressing cell, or in both.

25 KIM ligation-triggered response in a KIM ligand-expressing cell may be generated by contacting the cell with exogenous KIM, KIM fusion proteins or activating antibodies against KIM ligand, either in vitro or in vivo. Further, responses of the KIM ligand-expressing cell that would otherwise be triggered by endogenous KIM could be blocked by contacting the KIM ligand-expressing cell with a KIM ligand antagonist (e.g., an antagonist antibody that binds to

KIM ligand), or by contacting the endogenous KIM with an anti-KIM antibody or other KIM-binding molecule which prevents the effective ligation of KIM with a KIM ligand.

Similarly, the responses triggered by KIM ligation in the KIM-expressing cell may be promoted or inhibited with exogenous compounds. For example, KIM ligation-triggered  
5 response in a KIM-expressing cell may be generated by contacting the cell with a soluble KIM ligand, or certain anti-KIM activating antibodies. Further, responses of the KIM-expressing cell that would otherwise be triggered by interaction with endogenous KIM ligand could be blocked by contacting the KIM-expressing cell with an antagonist to KIM (e.g., a blocking antibody that binds to KIM in a manner that prevents effective, signal-generating KIM ligation), or by  
10 contacting the endogenous KIM ligand with an anti-KIM ligand antibody or other KIM ligand-binding molecule which prevents the effective ligation of KIM with the KIM ligand.

Which of the interventions described above are useful for particular therapeutic uses depend on the relevant etiologic mechanism of either the pathologic process to be inhibited, or of the medically desirable process to be promoted, as is apparent to those of skill in the medical  
15 arts. For example, where KIM ligation results in desirable cellular growth, maintenance of differentiated phenotype, resistance to apoptosis induced by various insults, or other medically advantageous responses, one of the above-described interventions that promote ligation-triggered response may be employed. In the alternative, one of the inhibitory interventions may be useful where KIM ligation invokes undesirable consequences, such as neoplastic growth, deleterious  
20 loss of cellular function, susceptibility to apoptosis, or promotion of inflammation events.

Following are examples of the previously described therapeutic methods of the invention. One therapeutic use of the KIM-related compounds of the invention is for treating a subject with renal disease, promoting growth of new tissue in a subject, or promoting survival of damaged tissue in a subject, and includes the step of administering to the subject a therapeutically  
25 effective amount of a KIM protein of the invention, or of a pharmaceutical composition which includes a protein of the invention. The protein used in these methods may be a fragment of a full-length KIM protein, a soluble KIM ligand protein or fusion fragment, or a KIM agonist. These methods may also be practiced by administering to the subject a therapeutically effective amount of an agonist antibody of the invention, or a pharmaceutical composition which includes  
30 an agonist antibody of the invention. A KIM protein may be administered concurrently with a therapeutically effective amount of a second compound which exerts a medically desirable

adjunct effect. While tissues of interest for these methods may include any tissue, preferred tissues include renal tissue, liver, neural tissue, heart, stomach, small intestine, spinal cord, or lung. Particular renal conditions which may be beneficially treated with the compounds of the invention include acute renal failure, acute nephritis, chronic renal failure, nephrotic syndrome, 5 renal tubule defects, kidney transplants, toxic injury, hypoxic injury, and trauma. Renal tubule defects include those of either hereditary or acquired nature, such as polycystic renal disease, medullary cystic disease, and medullary sponge kidney. This list is not limited, and may include many other renal disorders (see, e.g., Harrison's Principles of Internal Medicine, 13th ed., 1994, which is herein incorporated by reference.) The subject of the methods may be human.

10 A therapeutic intervention for inhibiting growth of undesirable, KIM ligand-expressing tissue in a subject includes the step of administering to the subject a therapeutically effective amount of a KIM antagonist (e.g., an antagonist antibody that binds to KIM ligand), or by administering a therapeutically effective amount of an anti-KIM antibody or other KIM-binding molecule which blocks KIM binding to the KIM ligand-expressing tissue. In an embodiment of 15 interest, the KIM antagonist or anti-KIM antibody may be used therapeutically to inhibit or block growth of tumors which depend on KIM protein for growth.

Other methods of the invention include killing KIM ligand-expressing tumor cells, or inhibiting their growth, by contacting the cells with a fusion protein of a KIM and a toxin or radionuclide, or an anti-KIM ligand antibody conjugated to a toxin or radionuclide. The cell may 20 be within a subject, and the protein or the conjugated antibody is administered to the subject.

Also encompassed within the invention is a method for targeting a toxin or radionuclide to a cell expressing a KIM, comprising contacting the cell with a fusion protein comprising a KIM ligand and a toxin or radionuclide, or an anti-KIM antibody conjugated to a toxin or radionuclide. Another embodiment includes the method of suppressing growth of a tumor cell 25 which expresses KIM, comprising contacting the cell with a fusion protein of KIM ligand and a toxin or radionuclide or with an anti-KIM antibody conjugated to a toxin or radionuclide; the cell may be within a subject, and the protein administered to the subject.

The term "subject" used herein is taken to mean any mammal to which KIM may be administered. Subjects specifically intended for treatment with the method of the invention 30 include humans, as well as nonhuman primates, sheep, horses, cattle, goats, pigs, dogs, cats,

- 25 -

rabbits, guinea pigs, hamsters, gerbils, rats and mice, as well as the organs, tumors, and cells derived or originating from these hosts.

#### Use of Compounds of the Invention in Gene Therapy

The KIM genes of the invention are introduced into damaged tissue, or into tissue where stimulated growth is desirable. Such gene therapy stimulates production of KIM protein by the transfected cells, promoting cell growth and/or survival of cells that express the KIM protein.

In a specific embodiment of a gene therapy method, a gene coding for a KIM protein may be introduced into a renal target tissue. The KIM protein would be stably expressed and stimulate tissue growth, division, or differentiation, or could potentiate cell survival.

Furthermore, a KIM gene may be introduced into a target cell using a variety of well-known methods that use either viral or non-viral based strategies.

Non-viral methods include electroporation, membrane fusion with liposomes, high velocity bombardment with DNA-coated microprojectiles, incubation with calcium-phosphate-DNA precipitate, DEAE-dextran mediated transfection, and direct micro-injection into single cells. For instance, a KIM gene may be introduced into a cell by calcium phosphate coprecipitation (Pillicer et al., *Science*, 209: 1414-1422 (1980); mechanical microinjection and/or particle acceleration (Anderson et al., *Proc. Nat. Acad. Sci. USA*, 77: 5399-5403 (1980); liposome based DNA transfer (e.g., LIPOFECTIN-mediated transfection- Fefgner et al., *Proc. Nat. Acad. Sci., USA*, 84: 471-477, 1987; Gao and Huang, *Biochim. Biophys. Res. Comm.*, 179: 280-285, 1991; DEAE Dextran-mediated transfection; electroporation (U.S. Patent 4,956,288); or polylysine-based methods in which DNA is conjugated to deliver DNA preferentially to liver hepatocytes (Wolff et al., *Science*, 247: 465-468, 1990; Curiel et al., *Human Gene Therapy* 3: 147-154, 1992).

Target cells may be transfected with the genes of the invention by direct gene transfer. See, e.g., Wolff et al., "Direct Gene Transfer Into Moose Muscle In Vivo", *Science* 247:1465-68, 1990. In many cases, vector-mediated transfection will be desirable. Any of the methods known in the art for the insertion of polynucleotide sequences into a vector may be used. ( See, for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; and Ausubel et al., *Current Protocols in Molecular Biology*, J. Wiley & Sons, NY, 1992, both of which are incorporated herein by reference.)

- 26 -

Promoter activation may be tissue specific or inducible by a metabolic product or administered substance. Such promoters/enhancers include, but are not limited to, the native c-ret ligand protein promoter, the cytomegalovirus immediate-early promoter/enhancer (Karasuyama et al., J. Exp. Med., 169: 13, 1989); the human beta-actin promoter (Gunning et al., Proc. Nat. Acad. Sci. USA, 84: 4831, 1987; the glucocorticoid-inducible promoter present in the mouse mammary tumor virus long terminal repeat (MMTV LTR) (Klessig et al., Mol. Cell. Biol., 4: 1354, 1984); the long terminal repeat sequences of Moloney murine leukemia virus (MuLV LTR) (Weiss et al., RNA Tumor Viruses, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1985); the SV40 early region promoter (Bernoist and Chambon, Nature, 290:304, 1981); the promoter of the Rous sarcoma virus (RSV) (Yamamoto et al., Cell, 22:787, 1980); the herpes simplex virus (HSV) thymidine kinase promoter (Wagner et al., Proc. Nat. Acad. Sci. USA, 78: 1441, 1981); the adenovirus promoter (Yamada et al., Proc. Nat. Acad. Sci. USA, 82: 3567, 1985).

The KIM genes may also be introduced by specific viral vectors for use in gene transfer systems which are now well established. See for example: Madzak et al., J. Gen . Virol., 73: 1533-36, 1992 (papovavirus SV40); Berkner et al., Curr. Top. Microbiol. Immunol., 158: 39-61, 1992 (adenovirus); Hofmann et al., Proc. Natl. Acad. Sci. 92: 10099-10103, 1995 (baculovirus); Moss et al., Curr. Top. Microbiol. Immunol., 158: 25-38, 1992 (vaccinia virus); Muzyczka, Curr. Top. Microbiol. Immunol., 158: 97-123, 1992 ( adeno-associated virus); Margulies, Curr. Top. Microbiol. Immunol., 158: 67-93, 1992 (herpes simplex virus (HSV) and Epstein-Barr virus (HBV)); Miller, Curr. Top. Microbiol. Immunol., 158: 1-24, 1992 (retrovirus); Brandyopadhyay et al.,Mol. Cell. Biol., 4: 749-754, 1984 (retrovirus); Miller et al., Nature, 357: 455-450, 1992 (retrovirus); Anderson, Science, 256: 808-813, 1992 (retrovirus), Current Protocols in Molecular Biology: Sections 9.10-9.14 (Ausubel et al., Eds.), Greene Publishing Associates, 1989, all of which are incorporated herein by reference.

Preferred vectors are DNA viruses that include adenoviruses (preferably Ad-2 or Ad-5 based vectors), baculovirus, herpes viruses (preferably herpes simplex virus based vectors), and parvoviruses (preferably "defective" or non-autonomous parvovirus based vectors, more preferably adeno-associated virus based vectors, most preferably AAV-2 based vectors). See, e.g., Ali et al., Gene Therapy 1: 367-384, 1994; U.S. Patent 4,797,368 and 5,399,346 and discussion below.

The choice of a particular vector system for transferring, for instance, a KIM sequence will depend on a variety of factors. One important factor is the nature of the target cell population. Although retroviral vectors have been extensively studied and used in a number of gene therapy applications, they are generally unsuited for infecting cells that are not dividing but may be  
5 useful in cancer therapy since they only integrate and express their genes in replicating cells. They are useful for *ex vivo* approaches and are attractive in this regard due to their stable integration into the target cell genome.

Adenoviruses are eukaryotic DNA viruses that can be modified to efficiently deliver a therapeutic or reporter transgene to a variety of cell types. The general adenoviruses types 2 and  
10 5 (Ad2 and Ad5, respectively), which cause respiratory disease in humans, are currently being developed for gene therapy of Duchenne Muscular Dystrophy (DMD) and Cystic Fibrosis (CF). Both Ad2 and Ad5 belong to a subclass of adenovirus that are not associated with human malignancies. Adenovirus vectors are capable of providing extremely high levels of transgene delivery to virtually all cell types, regardless of the mitotic state. High titers ( $10^{13}$  plaque  
15 forming units/ml) of recombinant virus can be easily generated in 293 cells (an adenovirus-transformed, complementation human embryonic kidney cell line: ATCC CRL1573) and cryo-stored for extended periods without appreciable losses. The efficacy of this system in delivering a therapeutic transgene *in vivo* that complements a genetic imbalance has been demonstrated in animal models of various disorders. See Watanabe, Atherosclerosis, 36: 261-268, 1986;  
20 Tanzawa et al., FEBS Letters 118(1):81-84, 1980; Golasten et al., New Engl.J. Med. 309:288-296, 1983; Ishibashi et al., J. Clin. Invest. 92: 883-893, 1993; and Ishibashi et al., J. Clin. Invest. 93: 1889-1893, 1994, all of which are incorporated herein by reference. Indeed, recombinant replication defective adenovirus encoding a cDNA for the cystic fibrosis transmembrane regulator (CFTR) has been approved for use in at least two human CF clinical trials. See, e.g.,  
25 Wilson, Nature 365:691-692, 1993. Further support of the safety of recombinant adenoviruses for gene therapy is the extensive experience of live adenovirus vaccines in human populations.

The first-generation recombinant, replication-deficient adenoviruses which have been developed for gene therapy of DMD and other inherited disorders contain deletions of the entire E1a and part of the E1b regions. This replication-defective virus is grown in 293 cells containing  
30 a functional adenovirus E1a gene which provides a transacting E1a protein. E1-deleted viruses are capable of replicating and producing infectious virus in the 293 cells, which provide E1a and

- 28 -

E1b region gene products in *trans*. The resulting virus is capable of infecting many cell types and can express the introduced gene (providing it carries its own promoter), but cannot replicate in a cell that does not carry the E1 region DNA unless the cell is infected at a very high multiplicity of infection. Adenoviruses have the advantage that they have a broad host range, can 5 infect quiescent or terminally differentiated cells such as neurons, and appear essentially non-oncogenic. Adenoviruses do not appear to integrate into the host genome. Because they exist extrachromosomally, the risk of insertional mutagenesis is greatly reduced. Ali et al., *supra*, at 373. Recombinant adenoviruses (rAdV) produce very high titers, the viral particles are moderately stable, expression levels are high, and a wide range of cells can be infected. Their 10 natural host cells are airway epithelium, so they are useful for therapy of lung cancers.

Baculovirus-mediated transfer has several advantages. Baculoviral gene transfer can occur in replicating and nonreplicating cells, and can occur in renal cells, as well as in hepatocytes, neural cells, spleen, skin, and muscle. Baculovirus is non-replicating and nonpathogenic in mammalian cells. Humans lack pre-existing antibodies to recombinant baculovirus which could 15 block infection. In addition, baculovirus is capable of incorporating and transducing very large DNA inserts.

Adeno-associated viruses (AAV) have also been employed as vectors for somatic gene therapy. AAV is a small, single-stranded (ss) DNA virus with a simple genomic organization (4-7 kb) that makes it an ideal substrate for genetic engineering. Two open reading frames encode a 20 series of *rep* and *cap* polypeptides. *Rep* polypeptides (*rep78*, *rep68*, *rep 62* and *rep 40*) are involved in replication, rescue and integration of the AAV genome. The *cap* proteins (VP1, VP2 and VP3) form the virion capsid. Flanking the *rep* and *cap* open reading frames at the 5' and 3' ends are 145 bp inverted terminal repeats (ITRs), the first 125 bp of which are capable of forming Y- or T-shaped duplex structures. Of importance for the development of AAV vectors, 25 the entire *rep* and *cap* domains can be excised and replaced with a therapeutic or reporter transgene. See B.J. Carter, in *Handbook of Parvoviruses*, ed., P. Tijsser, CRC Press, pp. 155-168 (1990). It has been shown that the ITRs represent the minimal sequence required for replication, rescue, packaging, and integration of the AAV genome.

Adeno-associated viruses (AAV) have significant potential in gene therapy. The viral 30 particles are very stable and recombinant AAVs (rAAV) have "drug-like" characteristics in that rAAV can be purified by pelleting or by CsCl gradient banding. They are heat stable and can be

lyophilized to a powder and rehydrated to full activity. Their DNA stably integrates into host chromosomes so expression is long-term. Their host range is broad and AAV causes no known disease so that the recombinant vectors are non-toxic.

Once introduced into a target cell, sequences of interest can be identified by conventional methods such as nucleic acid hybridization using probes comprising sequences that are homologous/complementary to the inserted gene sequences of the vector. In another approach, the sequence(s) may be identified by the presence or absence of a "marker" gene function (e.g., thymidine kinase activity, antibiotic resistance, and the like) caused by introduction of the expression vector into the target cell.

#### 10 Formulations and Administration

The compounds of the invention are formulated according to standard practice, such as prepared in a carrier vehicle. The term "pharmacologically acceptable carrier" means one or more organic or inorganic ingredients, natural or synthetic, with which the mutant proto-oncogene or mutant oncoprotein is combined to facilitate its application. A suitable carrier includes sterile saline although other aqueous and non-aqueous isotonic sterile solutions and sterile suspensions known to be pharmaceutically acceptable are known to those of ordinary skill in the art. In this regard, the term "carrier" encompasses liposomes and the HIV-1 tat protein (See Chen et al., Anal. Biochem. 227: 168-175, 1995) as well as any plasmid and viral expression vectors.

20 Any of the novel polypeptides of this invention may be used in the form of a pharmaceutically acceptable salt. Suitable acids and bases which are capable of forming salts with the polypeptides of the present invention are well known to those of skill in the art, and include inorganic and organic acids and bases.

A compound of the invention is administered to a subject in a therapeutically-effective amount, which means an amount of the compound which produces a medically desirable result or exerts an influence on the particular condition being treated. An effective amount of a compound of the invention is capable of ameliorating or delaying progression of the diseased, degenerative or damaged condition. The effective amount can be determined on an individual basis and will be based, in part, on consideration of the physical attributes of the subject,

- 30 -

symptoms to be treated and results sought. An effective amount can be determined by one of ordinary skill in the art employing such factors and using no more than routine experimentation.

A liposome delivery system for a compound of the invention may be any of a variety of unilamellar vesicles, multilamellar vesicles, or stable plurilamellar vesicles, and may be prepared 5 and administered according to methods well known to those of skill in the art, for example in accordance with the teachings of United States Patent 5,169,637, 4,762,915, 5,000,958 or 5,185,154. In addition, it may be desirable to express the novel polypeptides of this invention, as well as other selected polypeptides, as lipoproteins, in order to enhance their binding to liposomes. As an example, treatment of human acute renal failure with liposome-encapsulated 10 KIM protein may be performed in vivo by introducing a KIM protein into cells in need of such treatment using liposomes. The liposomes can be delivered via catheter to the renal artery. The recombinant KIM protein is purified, for example, from CHO cells by immunoaffinity chromatography or any other convenient method, then mixed with liposomes and incorporated 15 into them at high efficiency. The encapsulated protein may be tested in vitro for any effect on stimulating cell growth.

The compounds of the invention may be administered in any manner which is medically acceptable. This may include injections, by parenteral routes such as intravenous, intravascular, intraarterial, subcutaneous, intramuscular, intratumor, intraperitoneal, intraventricular, intraepidural, or others as well as oral, nasal, ophthalmic, rectal, or topical. Sustained release 20 administration is also specifically included in the invention, by such means as depot injections or erodible implants. Localized delivery is particularly contemplated, by such means as delivery via a catheter to one or more arteries, such as the renal artery or a vessel supplying a localized tumor.

Although the foregoing invention has been described in some detail by way of illustration 25 and example for purposes of clarity of understanding, it will be obvious to one skilled in the art that certain changes and modifications may be practiced within the scope of the invention, as limited only by the scope of the appended claims.

- 31 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Michele Sanicola-Nadel  
Joseph V. Bonventre  
Catherine A. Hession  
Takaharu Ichimura  
Henry Wei  
Richard L. Cate

(ii) TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Biogen, Inc.  
(B) STREET: 14 Cambridge Center  
(C) CITY: Cambridge  
(D) STATE: MA  
(E) COUNTRY: USA  
(F) ZIP: 02142

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE: 23-MAY-1997  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 60/018,228  
(B) FILING DATE: 24-MAY-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Levine, Leslie M.  
(B) REGISTRATION NUMBER: 35,245  
(C) REFERENCE/DOCKET NUMBER: A010 PCT CIP

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 679-2810  
(B) TELEFAX: (617) 679-2838

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2566 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

- 32 -

## (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 615..1535

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GC GG CG CG GT CG AC GG TG CC TGT GAG T AAA TAG ATC AGGG T CT C C TT CAC AGC AC AT T CT	60
CC AG GA AG CC GAG CAA AAC AT TAG TG CT ATT TT ACC CAG GA GG AA AT CT AG GT GT AG AG AG	120
CT CT AC GG AT CTA AGG TT TG GAT CT GT ACC CAG TG CT TT TT AGG TG CT TT AG AC AT TT	180
CT CAG GA AGA TG TAG T CT CT GTC ACC AT GT GT GG CT GA AT T CT AG CT CA G TCC AT CT TA	240
T GT GT TT AA AG GT AG TT GA AG TTT AG GA ACC A C CAG T AT G TCT CT GAG CA GA AG AGT AC A	300
G T GT CC AT CT TG AGG GACA AG CTC AT CT TT A CC AT TA GAG G GCT GG C CTT G GCT T AG AT TC	360
TAC CG AGA AAC AT ACT CT CTA AT GG CT GCCC TC AGT TT T CT CT G TTT G CT G TCT T AT TT GT	420
GTC AT GG CC A GA AGT C AT AT GG AT GG CT CT AT GT GAG CAA GG ACC CAG AT AGA AG AGT GT	480
A TTT GGG GGA AC AGG TT G CC CTA AC AGA GA GT C CT GT GGG AT TC AT G CAG TC AGG AT GAA	540
GAC CT GAT CA GAC AG AGT GT GCT GAG TG CC AC GG CT A ACC AG AGT GACT T GTC ACT GT CC	600
TTC AGG TCAA CACC ATG GTT CAA CTT CAA GTC TTC ATT TCA GGC CTC CTG Met Val Gln Leu Gln Val Phe Ile Ser Gly Leu Leu	650
1                   5                   10	
CT G CTT CTT CCA GGC TCT GTA GAT TCT TAT GAA GTA GT G AAG GGG GT G Leu Leu Pro Gly Ser Val Asp Ser Tyr Glu Val Val Lys Gly Val	698
15                 20                 25	
GT G GGT CAC CCT GTC ACA ATT CCA TGT ACT TAC TCA ACA CGT GGA GGA Val Gly His Pro Val Thr Ile Pro Cys Thr Tyr Ser Thr Arg Gly Gly	746
30                 35                 40	
AT C ACA ACG ACA TGT TGG GGC CGG GGG CAA TGC CCA TAT TCT AGT TGT Ile Thr Thr Cys Trp Gly Arg Gly Gln Cys Pro Tyr Ser Ser Cys	794
45                 50                 55                 60	
CAA AAT ATA CTT ATT TGG ACC AAT GGA TAC CAA GTC ACC TAT CGG AGC Gln Asn Ile Leu Ile Trp Thr Asn Gly Tyr Gln Val Thr Tyr Arg Ser	842
65                 70                 75	
AGC GGT CGA TAC AAC ATA AAG GGG CGT ATT TCA GAA GGA GAC GTA TCC Ser Gly Arg Tyr Asn Ile Lys Gly Arg Ile Ser Glu Gly Asp Val Ser	890
80                 85                 90	

- 33 -

TTG ACA ATA GAG AAC TCT GTT GAT AGT GAT AGT GGT CTG TAT TGT TGC Leu Thr Ile Glu Asn Ser Val Asp Ser Asp Ser Gly Leu Tyr Cys Cys 95 100 105	938
CGA GTG GAG ATT CCT GGA TGG TTC AAC GAT CAG AAA ATG ACC TTT TCA Arg Val Glu Ile Pro Gly Trp Phe Asn Asp Gln Lys Met Thr Phe Ser 110 115 120	986
TTG GAA GTT AAA CCA GAA ATT CCC ACA AGT CCT CCA ACA AGA CCC ACA Leu Glu Val Lys Pro Glu Ile Pro Thr Ser Pro Pro Thr Arg Pro Thr 125 130 135 140	1034
ACT ACA AGA CCC ACA ACC ACA AGG CCC ACA ACT ATT TCA ACA AGA TCC Thr Thr Arg Pro Thr Thr Arg Pro Thr Thr Ile Ser Thr Arg Ser 145 150 155	1082
ACA CAT GTA CCA ACA TCA ACC AGA GTC TCC ACC TCT ACT CCA ACA CCA Thr His Val Pro Thr Ser Thr Arg Val Ser Thr Ser Thr Pro Thr Pro 160 165 170	1130
GAA CAA ACA CAG ACT CAC AAA CCA GAA ATC ACT ACA TTT TAT GCC CAT Glu Gln Thr Gln Thr His Lys Pro Glu Ile Thr Thr Phe Tyr Ala His 175 180 185	1178
GAG ACA ACT GCT GAG GTG ACA GAA ACT CCA TCA TAT ACT CCT GCA GAC Glu Thr Thr Ala Glu Val Thr Glu Thr Pro Ser Tyr Thr Pro Ala Asp 190 195 200	1226
TGG AAT GGC ACT GTG ACA TCC TCA GAG GAG GCC TGG AAT AAT CAC ACT Trp Asn Gly Thr Val Thr Ser Ser Glu Glu Ala Trp Asn Asn His Thr 205 210 215 220	1274
GTA AGA ATC CCT TTG AGG AAG CCG CAG AGA AAC CCG ACT AAG GGC TTC Val Arg Ile Pro Leu Arg Lys Pro Gln Arg Asn Pro Thr Lys Gly Phe 225 230 235	1322
TAT GTT GGC ATG TCC GTT GCA GCC CTG CTG CTG CTG CTG CTT GCG AGC Tyr Val Gly Met Ser Val Ala Ala Leu Leu Leu Leu Leu Ala Ser 240 245 250	1370
ACC GTG GTT GTC ACC AGG TAC ATC ATT ATA AGA AAG AAG ATG GGC TCT Thr Val Val Val Thr Arg Tyr Ile Ile Arg Lys Lys Met Gly Ser 255 260 265	1418
CTG AGC TTT GTT GCC TTC CAT GTC TCT AAG AGT AGA GCT TTG CAG AAC Leu Ser Phe Val Ala Phe His Val Ser Lys Ser Arg Ala Leu Gln Asn 270 275 280	1466
GCA GCG ATT GTG CAT CCC CGA GCT GAA GAC AAC ATC TAC ATT ATT GAA Ala Ala Ile Val His Pro Arg Ala Glu Asp Asn Ile Tyr Ile Ile Glu 285 290 295 300	1514
GAT AGA TCT CGA GGT GCA GAA TGAGTCCCCAG AGGCCTTCTG TGGGGCCTTC Asp Arg Ser Arg Gly Ala Glu 305	1565

- 34 -

TGCCTGGGAT TACAGAGATC GTGACTGATT TCACAGAGTA AAATACCCAT TCCAGCTCCT	1625
GGGAGATTTT GTGTTTGTT TCTTCAGCT GCAGTGGAGA GGGTAACCCCT CTACCCCTGTA	1685
TATGCAAAAC TCGAGGTTAA CATCATCCTA ATTCTTGTAT CAGCAACACC TCAGTGTCTC	1745
CACTCACTGC AGCGATTCTC TCAAATGTGA ACATTTAGA AGTTTGTGTT TCCTTTGTC	1805
CATGTAATCA TTGGTAATAC AAGAATTATA TCTTGTATTATAAAAACCAATT AATGAGAGGG	1865
GAATAGGAAT TAAAAGCTGG TGGGAAGGGC CT CCTGAATT TAGAAGCACT TCATGATTGT	1925
GTTTATCTCT TTATTGTAA TTGAAATGT TACTTCTATC CTTCCAAGG GGCAAAATCA	1985
TGGGAGCATG GAGGTTTAA TTGCCCTCAT AGATAAGTAG AAGAAGAGAG TCTAATGCCA	2045
CCAATAGAGG TGTTTATGCT TTCTCACAGC TCTGGAAATA TGATCATTAA TTATGCAGTT	2105
GATCTTAGGA TGAGGATGGG TTTCTTAGGA GGAGAGGTTA CCATGGTGAG TGGACCAGGC	2165
ACACATCAGG GGAAGAAAAC AATGGATCAA GGGATTGAGT TCATTAGAGC CATTCCACT	2225
CCACTTCTGT CTTGATGTC AGTGTCCCTA AACTCACCCA CTGAGCTCTG AATTAGGTGC	2285
AGGGAGGAGA CGTGCAGAAA CGAAAGAGGA AAGAAAGGAG AGAGAGCAGG ACACAGGCTT	2345
TCTGCTGAGA GAAGTCCTAT TGCAGGTGTG ACAGTGTGTT GGACTACCAC GGGTTTCCTT	2405
CAGACTTCTA AGTTTCTAAA TCACTATCAT GTGATCATAT TTATTTTAA AATTATTTCA	2465
GAAAGACACC ACATTTCAA TAATAATCA GTTGTACA ATTAATAAAA TATTTGTTT	2525
GCTAAGAAGT AAAAAAAAAA AAAAAAGTC GACGCCCG C	2566

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2084 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 145..1065

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCGGCCGCGT CGACGGTGCC TGTGAGTAAA TAGATCAGGG TCTCCTTCAC AGCACATTCT

- 35 -

- 36 -

CCT GCA GAC TGG AAT GGC ACT GTG ACA TCC TCA GAG GAG GCC TGG AAT Pro Ala Asp Trp Asn Gly Thr Val Thr Ser Ser Glu Glu Ala Trp Asn 205 210 215	795
AAT CAC ACT GTA AGA ATC CCT TTG AGG AAG CCG CAG AGA AAC CCG ACT Asn His Thr Val Arg Ile Pro Leu Arg Lys Pro Gln Arg Asn Pro Thr 220 225 230	843
AAG GGC TTC TAT GTT GGC ATG TCC GTT GCA GCC CTG CTG CTG CTG CTG Lys Gly Phe Tyr Val Gly Met Ser Val Ala Ala Leu Leu Leu Leu 235 240 245	891
CTT GCG AGC ACC GTG GTT GTC ACC AGG TAC ATC ATT ATA AGA AAG AAG Leu Ala Ser Thr Val Val Val Thr Arg Tyr Ile Ile Ile Arg Lys Lys 250 255 260 265	939
ATG GGC TCT CTG AGC TTT GTT GCC TTC CAT GTC TCT AAG AGT AGA GCT Met Gly Ser Leu Ser Phe Val Ala Phe His Val Ser Lys Ser Arg Ala 270 275 280	987
TTG CAG AAC GCA GCG ATT GTG CAT CCC CGA GCT GAA GAC AAC ATC TAC Leu Gln Asn Ala Ala Ile Val His Pro Arg Ala Glu Asp Asn Ile Tyr 285 290 295	1035
ATT ATT GAA GAT AGA TCT CGA GGT GCA GAA TGAGTCCCAG AGGCCTTCTG Ile Ile Glu Asp Arg Ser Arg Gly Ala Glu 300 305	1085
TGGGGCCTTC TGCCTGGGAT TACAGAGATC GTGACTGATT TCACAGAGTA AAATACCCAT	1145
TCCAGCTCCT GGGAGATTTT GTGTTTGTT TCTTCCAGCT GCAGTGGAGA GGGTAACCCT	1205
CTACCCCTGTA TATGCAAAAC TCGAGGTTAA CATCATCCTA ATTCTTGAT CAGCAACACC	1265
TCAGTGTCTC CACTCACTGC AGCGATTCTC TCAAATGTGA ACATTTAGA AGTTTGTGTT	1325
TCCTTTGTC CATGTAATCA TTGGTAATAC AAGAATTATA TCTTGTATTAT TAAAACCATT	1385
AATGAGAGGG GAATAGGAAT TAAAAGCTGG TGGGAAGGGC CTCCTGAATT TAGAAGCACT	1445
TCATGATTGT GTTTATCTCT TTTATTGTAA TTTGAAATGT TACTTCTATC CTTCCCAAGG	1505
GGCAAAATCA TGGGAGCATG GAGGTTTAA TTGCCCTCAT AGATAAGTAG AAGAAGAGAG	1565
TCTAATGCCA CCAATAGAGG TGGTTATGCT TTCTCACAGC TCTGGAAATA TGATCATTAA	1625
TTATGCAGTT GATCTTAGGA TGAGGATGGG TTTCTTAGGA GGAGAGGTTA CCATGGTGAG	1685
TGGACCAGGC ACACATCAGG GGAAGAAAAC AATGGATCAA GGGATTGAGT TCATTAGAGC	1745
CATTTCCACT CCACCTCTGT CTTGATGCTC AGTGTTCCTA AACTCACCCA CTGAGCTCTG	1805
AATTAGGTGC AGGGAGGAGA CGTGCAGAAA CGAAAGAGGA AAGAAAGGAG AGAGAGCAGG	1865
ACACAGGCTT TCTGCTGAGA GAAGTCCTAT TGCAGGTGTG ACAGTGTGTTG GGACTACCAC	1925

- 37 -

GGGTTTCCTT CAGACTTCTA AGTTTCTAAA TCACTATCAT GTGATCATAT TTATTTTTAA	1985
AATTATTTCA GAAAGACACC ACATTTCAA TAATAATCA GTTTGTACA ATTAATAAAA	2045
TATTTTGTAA GCTAAGAAGT AAAAAGTCGA CGCGGCCGC	2084

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 307 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Val Gln Leu Gln Val Phe Ile Ser Gly Leu Leu Leu Leu Pro			
1	5	10	15
Gly Ser Val Asp Ser Tyr Glu Val Val Lys Gly Val Val Gly His Pro			
20	25	30	
Val Thr Ile Pro Cys Thr Tyr Ser Thr Arg Gly Gly Ile Thr Thr Thr			
35	40	45	
Cys Trp Gly Arg Gly Gln Cys Pro Tyr Ser Ser Cys Gln Asn Ile Leu			
50	55	60	
Ile Trp Thr Asn Gly Tyr Gln Val Thr Tyr Arg Ser Ser Gly Arg Tyr			
65	70	75	80
Asn Ile Lys Gly Arg Ile Ser Glu Gly Asp Val Ser Leu Thr Ile Glu			
85	90	95	
Asn Ser Val Asp Ser Asp Ser Gly Leu Tyr Cys Cys Arg Val Glu Ile			
100	105	110	
Pro Gly Trp Phe Asn Asp Gln Lys Met Thr Phe Ser Leu Glu Val Lys			
115	120	125	
Pro Glu Ile Pro Thr Ser Pro Pro Thr Arg Pro Thr Thr Thr Arg Pro			
130	135	140	
Thr Thr Thr Arg Pro Thr Thr Ile Ser Thr Arg Ser Thr His Val Pro			
145	150	155	160
Thr Ser Thr Arg Val Ser Thr Ser Thr Pro Thr Pro Glu Gln Thr Gln			
165	170	175	
Thr His Lys Pro Glu Ile Thr Thr Phe Tyr Ala His Glu Thr Thr Ala			
180	185	190	

- 38 -

Glu Val Thr Glu Thr Pro Ser Tyr Thr Pro Ala Asp Trp Asn Gly Thr  
 195                    200                    205

Leu Arg Lys Pro Gln Arg Asn Pro Thr Lys Gly Phe Tyr Val Gly Met  
225 230 235 240

Ser Val Ala Ala Leu Leu Leu Leu Leu Ala Ser Thr Val Val Val  
245 250 255

Thr Arg Tyr Ile Ile Arg Lys Lys Met Gly Ser Leu Ser Phe Val  
260 265 270

Ala Phe His Val Ser Lys Ser Arg Ala Leu Gln Asn Ala Ala Ile Val  
275 280 285

His Pro Arg Ala Glu Asp Asn Ile Tyr Ile Ile Glu Asp Arg Ser Arg  
290 295 300

Gly Ala Glu  
305

(2) INFORMATION FOR SEQ ID NO:4:

**(i) SEQUENCE CHARACTERISTICS:**

- (A) LENGTH: 2303 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- FEATURE:**  
(a) MOUNT (SHELL) GPO

TCAGAGTCAA GCCCTGACTG GCTGAGGGCG GGCGCTCCGA GTCAGC ATG GAA AGT 115  
Met Glu Ser

CTC TGC GGG GTC CTG GTA TTT CTG CTG CTG GCT GCA GGA CTG CCG CTC 163  
 Leu Cys Gly Val Leu Val Phe Leu Leu Leu Ala Ala Gly Leu Pro Leu  
       5                 10                 15

CAG GCG GCC AAG CGG TTC CGT GAT GTG CTG GGC CAT GAG CAG TAT CCG      211  
 Gln Ala Ala Lys Arg Phe Arg Asp Val Leu Gly His Glu Gln Tyr Pro  
   20                25                30                35

- 39 -

GAT CAC ATG AGG GAG AAC AAC CAA TTA CGT GGC TGG TCT TCA GAT GAA Asp His Met Arg Glu Asn Asn Gln Leu Arg Gly Trp Ser Ser Asp Glu 40 45 50	259
AAT GAA TGG GAT GAA CAG CTG TAT CCA GTG TGG AGG AGG GGA GAG GGC Asn Glu Trp Asp Glu Gln Leu Tyr Pro Val Trp Arg Arg Gly Glu Gly 55 60 65	307
AGA TGG AAG GAC TCC TGG GAA GGA GGC CGT GTG CAG GCA GCC CTA ACC Arg Trp Lys Asp Ser Trp Glu Gly Gly Arg Val Gln Ala Ala Leu Thr 70 75 80	355
AGT GAT TCA CCG GCC TTG GTG GGT TCC AAT ATC ACC TTC GTA GTG AAC Ser Asp Ser Pro Ala Leu Val Gly Ser Asn Ile Thr Phe Val Val Asn 85 90 95	403
CTG GTG TTC CCC AGA TGC CAG AAG GAA GAT GCC AAC GGC AAT ATC GTC Leu Val Phe Pro Arg Cys Gln Lys Glu Asp Ala Asn Gly Asn Ile Val 100 105 110 115	451
TAT GAG AGG AAC TGC AGA AGT GAT TTG GAG CTG GCT TCT GAC CCG TAT Tyr Glu Arg Asn Cys Arg Ser Asp Leu Glu Leu Ala Ser Asp Pro Tyr 120 125 130	499
GTC TAC AAC TGG ACC ACA GGG GCA GAC GAT GAG GAC TGG GAA GAC AGC Val Tyr Asn Trp Thr Thr Gly Ala Asp Asp Trp Glu Asp Ser 135 140 145	547
ACC AGC CAA GGC CAG CAC CTC AGG TTC CCC GAC GGG AAG CCC TTC CCT Thr Ser Gln Gly Gln His Leu Arg Phe Pro Asp Gly Lys Pro Phe Pro 150 155 160	595
CGC CCC CAC GGA CGG AAG AAA TGG AAC TTC GTC TAC GTC TTC CAC ACA Arg Pro His Gly Arg Lys Lys Trp Asn Phe Val Tyr Val Phe His Thr 165 170 175	643
CTT GGT CAG TAT TTT CAA AAG CTG GGT CGG TGT TCA GCA CGA GTT TCT Leu Gly Gln Tyr Phe Gln Lys Leu Gly Arg Cys Ser Ala Arg Val Ser 180 185 190 195	691
ATA AAC ACA GTC AAC TTG ACA GTT GGC CCT CAG GTC ATG GAA GTG ATT Ile Asn Thr Val Asn Leu Thr Val Gly Pro Gln Val Met Glu Val Ile 200 205 210	739
GTC TTT CGA AGA CAC GGC CGG GCA TAC ATT CCC ATC TCC AAA GTG AAA Val Phe Arg Arg His Gly Arg Ala Tyr Ile Pro Ile Ser Lys Val Lys 215 220 225	787
GAC GTG TAT GTG ATA ACA GAT CAG ATC CCT ATA TTC GTG ACC ATG TAC Asp Val Tyr Val Ile Thr Asp Gln Ile Pro Ile Phe Val Thr Met Tyr 230 235 240	835
CAG AAG AAT GAC CGG AAC TCG TCT GAT GAA ACC TTC CTC AGA GAC CTC Gln Lys Asn Asp Arg Asn Ser Ser Asp Glu Thr Phe Leu Arg Asp Leu 245 250 255	883

- 40 -

CCC ATT TTC TTC GAT GTC CTC ATT CAC GAT CCC AGT CAT TTC CTC AAC Pro Ile Phe Phe Asp Val Leu Ile His Asp Pro Ser His Phe Leu Asn 260 265 270 275	931
TAC TCT GCC ATT TCC TAC AAG TGG AAC TTT GGG GAC AAC ACT GGC CTG Tyr Ser Ala Ile Ser Tyr Lys Trp Asn Phe Gly Asp Asn Thr Gly Leu 280 285 290	979
TTT GTC TCC AAC AAT CAC ACT TTG AAT CAC ACG TAT GTG CTC AAT GGA Phe Val Ser Asn Asn His Thr Leu Asn His Thr Tyr Val Leu Asn Gly 295 300 305	1027
ACC TTC AAC TTT AAC CTC ACC GTG CAA ACT GCA GTG CCG GGA CCA TGC Thr Phe Asn Phe Asn Leu Thr Val Gln Thr Ala Val Pro Gly Pro Cys 310 315 320	1075
CCC TCA CCC ACA CCT TCG CCT TCT TCT TCG ACT TCT CCT TCG CCT GCA Pro Ser Pro Thr Pro Ser Ser Ser Thr Ser Pro Ser Pro Ala 325 330 335	1123
TCT TCG CCT TCA CCC ACA TTA TCA ACA CCT AGT CCC TCT TTA ATG CCT Ser Ser Pro Ser Pro Thr Leu Ser Thr Pro Ser Pro Ser Leu Met Pro 340 345 350 355	1171
ACT GGC CAC AAA TCC ATG GAG CTG AGT GAC ATT TCC AAT GAA AAC TGC Thr Gly His Lys Ser Met Glu Leu Ser Asp Ile Ser Asn Glu Asn Cys 360 365 370	1219
CGA ATA AAC AGA TAT GGT TAC TTC AGA GCC ACC ATC ACA ATT GTA GAT Arg Ile Asn Arg Tyr Gly Tyr Phe Arg Ala Thr Ile Thr Ile Val Asp 375 380 385	1267
GGA ATC CTA GAA GTC AAC ATC ATC CAG GTA GCA GAT GTC CCA ATC CCC Gly Ile Leu Glu Val Asn Ile Ile Gln Val Ala Asp Val Pro Ile Pro 390 395 400	1315
ACA CCG CAG CCT GAC AAC TCA CTG ATG GAC TTC ATT GTG ACC TGC AAA Thr Pro Gln Pro Asp Asn Ser Leu Met Asp Phe Ile Val Thr Cys Lys 405 410 415	1363
GGG GCC ACT CCC ACG GAA GCC TGT ACG ATC ATC TCT GAC CCC ACC TGC Gly Ala Thr Pro Thr Glu Ala Cys Thr Ile Ile Ser Asp Pro Thr Cys 420 425 430 435	1411
CAG ATC GCC CAG AAC AGG GTG TGC AGC CCG GTG GCT GTG GAT GAG CTG Gln Ile Ala Gln Asn Arg Val Cys Ser Pro Val Ala Val Asp Glu Leu 440 445 450	1459
TGC CTC CTG TCC GTG AGG AGA GCC TTC AAT GGG TCC GGC ACG TAC TGT Cys Leu Leu Ser Val Arg Arg Ala Phe Asn Gly Ser Gly Thr Tyr Cys 455 460 465	1507
GTG AAT TTC ACT CTG GGA GAC GAT GCA AGC CTG GCC CTC ACC AGC GCC Val Asn Phe Thr Leu Gly Asp Asp Ala Ser Leu Ala Leu Thr Ser Ala 470 475 480	1555

- 41 -

CTG ATC TCT ATC CCT GGC AAA GAC CTA GGC TCC CCT CTG AGA ACA GTG Leu Ile Ser Ile Pro Gly Lys Asp Leu Gly Ser Pro Leu Arg Thr Val 485 490 495	1603
AAT GGT GTC CTG ATC TCC ATT GGC TGC CTG GCC ATG TTT GTC ACC ATG Asn Gly Val Leu Ile Ser Ile Gly Cys Leu Ala Met Phe Val Thr Met 500 505 510 515	1651
GTT ACC ATC TTG CTG TAC AAA AAA CAC AAG ACG TAC AAG CCA ATA GGA Val Thr Ile Leu Leu Tyr Lys Lys His Lys Thr Tyr Lys Pro Ile Gly 520 525 530	1699
AAC TGC ACC AGG AAC GTG GTC AAG GGC AAA GGC CTG AGT GTT TTT CTC Asn Cys Thr Arg Asn Val Val Lys Gly Lys Gly Leu Ser Val Phe Leu 535 540 545	1747
AGC CAT GCA AAA GCC CCG TTC TCC CGA GGA GAC CGG GAG AAG GAT CCA Ser His Ala Lys Ala Pro Phe Ser Arg Gly Asp Arg Glu Lys Asp Pro 550 555 560	1795
CTG CTC CAG GAC AAG CCA TGG ATG CTC TAAGTCTTCA CTCTCACTTC Leu Leu Gln Asp Lys Pro Trp Met Leu 565 570	1842
TGACTGGAA CCCACTCTTC TGTGCATGTA TGTGAGCTGT GCAGAAGTAC ATGACTGGTA GCTGTTGTTT TCTACGGATT ATTGTAAAAT GTATATCATG GTTAGGGAG CGTAGTTAAT	1902
TGGCATTTTA GTGAAGGGAT GGGAAAGACAG TATTTCTTCA CATCTGTATT GTGGTTTTA TACTGTTAAT AGGGTGGGCA CATTGTGTCT GAAGGGGGAG GGGGAGGTCA CTGCTACTTA	1962
AGGTCTAGG TTAACTGGGA GAGGATGCC CAGGCTCCTT AGATTTCTAC ACAAGATGTG CCTGAACCCA GCTAGTCCTG ACCTAAAGGC CATGCTTCAT CAACTCTATC TCAGCTCATT	2022
GAACATACCT GAGCACCTGA TGGAATTATA ATGGAACCAA GCTTGTGTA TGGTGTGTGT GTGTACATAA GATACTCATT AAAAGACAG TCTATTAAAA A	2082
	2142
	2202
	2262
	2303

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 572 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Glu Ser Leu Cys Gly Val Leu Val Phe Leu Leu Leu Ala Ala Gly  
1 5 10 15

- 42 -

Leu Pro Leu Gln Ala Ala Lys Arg Phe Arg Asp Val Leu Gly His Glu  
20 25 30

Gln Tyr Pro Asp His Met Arg Glu Asn Asn Gln Leu Arg Gly Trp Ser  
35 40 45

Ser Asp Glu Asn Glu Trp Asp Glu Gln Leu Tyr Pro Val Trp Arg Arg  
50 55 60

Gly Glu Gly Arg Trp Lys Asp Ser Trp Glu Gly Gly Arg Val Gln Ala  
65 70 75 80

Ala Leu Thr Ser Asp Ser Pro Ala Leu Val Gly Ser Asn Ile Thr Phe  
85 90 95

Val Val Asn Leu Val Phe Pro Arg Cys Gln Lys Glu Asp Ala Asn Gly  
100 105 110

Asn Ile Val Tyr Glu Arg Asn Cys Arg Ser Asp Leu Glu Leu Ala Ser  
115 120 125

Asp Pro Tyr Val Tyr Asn Trp Thr Thr Gly Ala Asp Asp Glu Asp Trp  
130 135 140

Glu Asp Ser Thr Ser Gln Gly Gln His Leu Arg Phe Pro Asp Gly Lys  
145 150 155 160

Pro Phe Pro Arg Pro His Gly Arg Lys Lys Trp Asn Phe Val Tyr Val  
165 170 175

Phe His Thr Leu Gly Gln Tyr Phe Gln Lys Leu Gly Arg Cys Ser Ala  
180 185 190

Arg Val Ser Ile Asn Thr Val Asn Leu Thr Val Gly Pro Gln Val Met  
195 200 205

Glu Val Ile Val Phe Arg Arg His Gly Arg Ala Tyr Ile Pro Ile Ser  
210 215 220

Lys Val Lys Asp Val Tyr Val Ile Thr Asp Gln Ile Pro Ile Phe Val  
225 230 235 240

Thr Met Tyr Gln Lys Asn Asp Arg Asn Ser Ser Asp Glu Thr Phe Leu  
245 250 255

Arg Asp Leu Pro Ile Phe Phe Asp Val Leu Ile His Asp Pro Ser His  
260 265 270

Phe Leu Asn Tyr Ser Ala Ile Ser Tyr Lys Trp Asn Phe Gly Asp Asn  
275 280 285

Thr Gly Leu Phe Val Ser Asn Asn His Thr Leu Asn His Thr Tyr Val  
290 295 300

- 43 -

Leu Asn Gly Thr Phe Asn Phe Asn Leu Thr Val Gln Thr Ala Val Pro			
305	310	315	320
Gly Pro Cys Pro Ser Pro Thr Pro Ser Pro Ser Ser Ser Thr Ser Pro			
325	330	335	
Ser Pro Ala Ser Ser Pro Ser Pro Thr Leu Ser Thr Pro Ser Pro Ser			
340	345	350	
Leu Met Pro Thr Gly His Lys Ser Met Glu Leu Ser Asp Ile Ser Asn			
355	360	365	
Glu Asn Cys Arg Ile Asn Arg Tyr Gly Tyr Phe Arg Ala Thr Ile Thr			
370	375	380	
Ile Val Asp Gly Ile Leu Glu Val Asn Ile Ile Gln Val Ala Asp Val			
385	390	395	400
Pro Ile Pro Thr Pro Gln Pro Asp Asn Ser Leu Met Asp Phe Ile Val			
405	410	415	
Thr Cys Lys Gly Ala Thr Pro Thr Glu Ala Cys Thr Ile Ile Ser Asp			
420	425	430	
Pro Thr Cys Gln Ile Ala Gln Asn Arg Val Cys Ser Pro Val Ala Val			
435	440	445	
Asp Glu Leu Cys Leu Leu Ser Val Arg Arg Ala Phe Asn Gly Ser Gly			
450	455	460	
Thr Tyr Cys Val Asn Phe Thr Leu Gly Asp Asp Ala Ser Leu Ala Leu			
465	470	475	480
Thr Ser Ala Leu Ile Ser Ile Pro Gly Lys Asp Leu Gly Ser Pro Leu			
485	490	495	
Arg Thr Val Asn Gly Val Leu Ile Ser Ile Gly Cys Leu Ala Met Phe			
500	505	510	
Val Thr Met Val Thr Ile Leu Leu Tyr Lys Lys His Lys Thr Tyr Lys			
515	520	525	
Pro Ile Gly Asn Cys Thr Arg Asn Val Val Lys Gly Lys Gly Leu Ser			
530	535	540	
Val Phe Leu Ser His Ala Lys Ala Pro Phe Ser Arg Gly Asp Arg Glu			
545	550	555	560
Lys Asp Pro Leu Leu Gln Asp Lys Pro Trp Met Leu			
565	570		

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1795 base pairs

- 44 -

- (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 278..1279

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

- 45 -

ACC GTA TCA TTG GAG ATT GTG CCA CCC AAG GTC ACG ACT ACT CCA ATT Thr Val Ser Leu Glu Ile Val Pro Pro Lys Val Thr Thr Thr Pro Ile 120 125 130	679
GTC ACA ACT GTT CCA ACC GTC ACG ACT GTT CGA ACG AGC ACC ACT GTT Val Thr Thr Val Pro Thr Val Thr Val Arg Thr Ser Thr Thr Val 135 140 145 150	727
CCA ACG ACA ACG ACT GTT CCA ACG ACA ACT GTT CCA ACA ACA ATG AGC Pro Thr Thr Thr Val Pro Thr Thr Val Pro Thr Thr Met Ser 155 160 165	775
ATT CCA ACG ACA ACG ACT GTT CCG ACG ACA ATG ACT GTT TCA ACG ACA Ile Pro Thr Thr Thr Val Pro Thr Thr Met Thr Val Ser Thr Thr 170 175 180	823
ACG AGC GTT CCA ACG ACA ACG AGC ATT CCA ACA ACA ACA AGT GTT CCA Thr Ser Val Pro Thr Thr Ser Ile Pro Thr Thr Thr Ser Val Pro 185 190 195	871
GTC ACA ACA ACG GTC TCT ACC TTT GTT CCT CCA ATG CCT TTG CCC AGG Val Thr Thr Thr Val Ser Thr Phe Val Pro Pro Met Pro Leu Pro Arg 200 205 210	919
CAG AAC CAT GAA CCA GTA GCC ACT TCA CCA TCT TCA CCT CAG CCA GCA Gln Asn His Glu Pro Val Ala Thr Ser Pro Ser Ser Pro Gln Pro Ala 215 220 225 230	967
GAA ACC CAC CCT ACG ACA CTG CAG GGA GCA ATA AGG AGA GAA CCC ACC Glu Thr His Pro Thr Thr Leu Gln Gly Ala Ile Arg Arg Glu Pro Thr 235 240 245	1015
AGC TCA CCA TTG TAC TCT TAC ACA ACA GAT GGG AAT GAC ACC GTG ACA Ser Ser Pro Leu Tyr Ser Tyr Thr Thr Asp Gly Asn Asp Thr Val Thr 250 255 260	1063
GAG TCT TCA GAT GGC CTT TGG AAT AAC AAT CAA ACT CAA CTG TTC CTA Glu Ser Ser Asp Gly Leu Trp Asn Asn Gln Thr Gln Leu Phe Leu 265 270 275	1111
GAA CAT AGT CTA CTG ACG GCC AAT ACC ACT AAA GGA ATC TAT GCT GGA Glu His Ser Leu Leu Thr Ala Asn Thr Thr Lys Gly Ile Tyr Ala Gly 280 285 290	1159
GTC TGT ATT TCT GTC TTG GTG CTT CTT GCT CTT TTG GGT GTC ATC ATT Val Cys Ile Ser Val Leu Val Leu Ala Leu Leu Gly Val Ile Ile 295 300 305 310	1207
GCC AAA AAG TAT TTC TTC AAA AAG GAG GTT CAA CAA CTA AGA CCC CAT Ala Lys Lys Tyr Phe Phe Lys Lys Glu Val Gln Gln Leu Arg Pro His 315 320 325	1255
AAA TCC TGT ATA CAT CAA AGA GAA TAGCCCTGG AAACATAGCA AATGAACTTC Lys Ser Cys Ile His Gln Arg Glu 330	1309

- 46 -

TATCTTGGCC ATCACAGCTG TCCAGAAGAG GGGAAATCTGT CTTAAAAACC AGCAAATCCA	1369
ACGTGAGACT TCATTTGGAA GCATTGTATG ATTATCTCTT GTTTCTATGT TATACTTCCA	1429
AATGTTGCAT TTCCATATGTT TTCCAAAGGT TTCAAATCGT GGGTTTTAT TTCCCTCCGTG	1489
GGGAAACAAA GTGAGTCTAA CTCACAGGTT TAGCTGTTT CTCATAACTC TGGAAATGTG	1549
ATGCATTAAG TACTGGATCT CTGAATTGGG GTAGCTGTTT TACCAAGTAA AGAGCCTACA	1609
ATAGTATGGA ACACATAGAC ACCAGGGAA GAAAATCATT TGCCAGGTGA TTTAACATAT	1669
TTATGCAATT TTTTTTTTTT TTTTGAGAT GGAGCTTGC TCTTGTGCC CAGGCTGGAG	1729
TGCGATGGTG AAATCTCGGC TCACTGTAAC CTCCACCTTC CGGGTTCAAG CAATTCTCCC	1789
GTCGAC	1795

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 334 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met His Pro Gln Val Val Ile Leu Ser Leu Ile Leu His Leu Ala Asp			
1	5	10	15
Ser Val Ala Gly Ser Val Lys Val Gly Gly Glu Ala Gly Pro Ser Val			
20	25	30	
Thr Leu Pro Cys His Tyr Ser Gly Ala Val Thr Ser Met Cys Trp Asn			
35	40	45	
Arg Gly Ser Cys Ser Leu Phe Thr Cys Gln Asn Gly Ile Val Trp Thr			
50	55	60	
Asn Gly Thr His Val Thr Tyr Arg Lys Asp Thr Arg Tyr Lys Leu Leu			
65	70	75	80
Gly Asp Leu Ser Arg Arg Asp Val Ser Leu Thr Ile Glu Asn Thr Ala			
85	90	95	
Val Ser Asp Ser Gly Val Tyr Cys Cys Arg Val Glu His Arg Gly Trp			
100	105	110	
Phe Asn Asp Met Lys Ile Thr Val Ser Leu Glu Ile Val Pro Pro Lys			
115	120	125	

- 47 -

Val Thr Thr Thr Pro Ile Val Thr Thr Val Pro Thr Val Thr Thr Val  
130 135 140

Arg Thr Ser Thr Thr Val Pro Thr Thr Thr Val Pro Thr Thr Thr  
145 150 155 160

Val Pro Thr Thr Met Ser Ile Pro Thr Thr Thr Val Pro Thr Thr  
165 170 175

Met Thr Val Ser Thr Thr Ser Val Pro Thr Thr Thr Ser Ile Pro  
180 185 190

Thr Thr Thr Ser Val Pro Val Thr Thr Val Ser Thr Phe Val Pro  
195 200 205

Pro Met Pro Leu Pro Arg Gln Asn His Glu Pro Val Ala Thr Ser Pro  
210 215 220

Ser Ser Pro Gln Pro Ala Glu Thr His Pro Thr Thr Leu Gln Gly Ala  
225 230 235 240

Ile Arg Arg Glu Pro Thr Ser Ser Pro Leu Tyr Ser Tyr Thr Thr Asp  
245 250 255

Gly Asn Asp Thr Val Thr Glu Ser Ser Asp Gly Leu Trp Asn Asn Asn  
260 265 270

Gln Thr Gln Leu Phe Leu Glu His Ser Leu Leu Thr Ala Asn Thr Thr  
275 280 285

Lys Gly Ile Tyr Ala Gly Val Cys Ile Ser Val Leu Val Leu Ala  
290 295 300

Leu Leu Gly Val Ile Ile Ala Lys Lys Tyr Phe Phe Lys Lys Glu Val  
305 310 315 320

Gln Gln Leu Arg Pro His Lys Ser Cys Ile His Gln Arg Glu  
325 330

- 48 -

What is claimed is:

1        1. A purified and isolated DNA molecule having a nucleotide sequence set forth in SEQ ID  
2        NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6.

1        2. A purified and isolated DNA molecule selected from among:  
2            a) the DNA molecule of SEQ ID NO:1 or its complementary strand;  
3            b) the DNA molecule of SEQ ID NO:2 or its complementary strand;  
4            c) the DNA molecule of SEQ ID NO:4 or its complementary strand;  
5            d) the DNA molecule of SEQ ID NO:6 or its complementary strand;  
6            e) DNA molecules which hybridize under stringent conditions to the DNA molecule  
7            defined in a), b), c) or d), or fragments thereof;  
8            f) DNA molecules which, but for the degeneracy of the genetic code, would hybridize to  
9            the DNA molecule defined in a), b), c), d) or e).

1        3. The recombinant DNA molecule according to claim 1 or 2, operably linked to an  
2        expression control sequence.

1        4. A vector comprising a purified and isolated DNA molecule having a nucleotide sequence  
2        set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6.

1        5. A biologically functional plasmid or viral DNA vector comprising a DNA molecule  
2        according to one of claims 1, 2 or 3.

1        6. A prokaryotic or eukaryotic host cell stably transformed or transfected by a vector  
2        comprising a DNA molecule of claim 1.

1        7. A process for the production of a polypeptide product encoded by a DNA molecule  
2        according to claim 1, 2 or 3, said process comprising:

- 49 -

3 growing, under suitable culture conditions, prokaryotic or eukaryotic host cells transformed  
4 or transfected with the DNA molecule in a manner allowing expression of the DNA  
5 molecule, and recovering the polypeptide product of said expression.

1 8. A polypeptide product produced by the process of claim 7.

1 9. A protein having an amino acid sequence which comprises SEQ ID NO:3, SEQ ID NO:5  
2 or SEQ ID NO:7.

1 10. A purified and isolated protein encoded by the DNA of SEQ ID NO:1, SEQ ID NO:2,  
2 SEQ ID NO:4 or SEQ ID NO:6.

1 11. The protein of claim 9 or 10, substantially free of other human proteins.

1 12. A protein which is a variant of SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:7.

1 13. A soluble variant of the protein according to claim 9, 10, 11 or 12.

1 14. An IgG fusion protein comprising the protein of claim 9, 10, 11, 12 or 13.

1 15. The soluble protein of claim 13, fused to a toxin, imageable compound or radionuclide.

1 16. A specific monoclonal antibody to a protein of claim 9, 10, 11 or 12.

1 17. The antibody of claim 16, associated with a toxin, imageable compound or radionuclide.

1 18. A hybridoma cell line which produces a specific antibody to the protein of claim 9, 10,  
2 11, 12 or 13.

1 19. An antibody produced by a hybridoma of claim 18.

- 50 -

1        20. A pharmaceutical composition comprising a therapeutically effective amount of the  
2 protein of claim 9, 10, 11, 12, 13, 14 or 15, and further comprising a pharmacologically  
3 acceptable carrier.

1        21. A pharmaceutical composition comprising a therapeutically effective amount of the  
2 antibody of claim 16, 17 or 19, and further comprising a pharmacologically acceptable carrier.

1        22. A method of treating a subject with renal disease, comprising administering to the  
2 subject a therapeutically effective amount of the protein of claim 9, 10, 11, 12, 13, 14 or 15.

1        23. A method of treating a subject with renal disease, comprising administering to the  
2 subject a therapeutically effective amount of the antibody of claim 16, 17 or 19.

1        24. A method of treating a subject with renal disease, comprising administering to the  
2 subject a therapeutically effective amount of the pharmaceutical composition of claim 20.

1        25. A method of promoting growth of new tissue in a subject, comprising administering to  
2 the subject a therapeutically effective amount of the protein of claim 9, 10, 11, 12, 13 or 14.

1        26. The method of claim 25, wherein the tissue is renal tissue.

1        27. A method of promoting survival of damaged tissue in a subject, comprising  
2 administering to the subject a therapeutically effective amount of the protein of claim 9, 10, 11,  
3 12, 13 or 14.

1        28. The method of claim 27, wherein the tissue is renal tissue.

1        29. A method of treating a subject with renal disease, comprising administering to the  
2 subject a therapeutically effective amount of the antibody of claim 16, 17 or 19.

- 51 -

1        30. A method of treating a subject with renal disease, comprising administering to the  
2        subject a therapeutically effective amount of the pharmaceutical composition of claim 21.

1        31. A method of promoting growth of new tissue in a subject, comprising administering to  
2        the subject a therapeutically effective amount of the antibody of claim 16, 17 or 19.

1        32. A method of promoting survival of damaged tissue in a subject, comprising  
2        administering to the subject a therapeutically effective amount of the antibody of claim 16, 17 or  
3        19.

1        33. A method of treating a subject with a renal disorder, comprising administering to the  
2        subject a vector of claim 4 or 5.

1        34. A method of promoting growth of new tissue in a subject, comprising administering to  
2        the subject a vector of claim 4 or 5.

1        35. A method of promoting survival of damaged tissue in a subject, comprising  
2        administering a therapeutically effective amount of a vector of claim 4 or 5 to the subject.

1        36. The method of claim 34 or 35, wherein the tissue is renal tissue.

1        37. A method for targeting an imageable compound to a cell expressing a protein of SEQ  
2        ID NO:3, SEQ ID NO:5 or SEQ ID NO:7, comprising contacting the cell with a monoclonal  
3        antibody of claim 16 fused to an imageable compound.

1        38. The method of claim 37, wherein the cell is within a subject, and the monoclonal  
2        antibody is administered to the subject.

1        39. A method of identifying damage or regeneration of renal cells in a subject, comprising  
2        comparing level of expression of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6

- 52 -

3    in renal cells of the subject to a control level of expression of SEQ ID NO:1, SEQ ID NO:2, SEQ  
4    ID NO:4 or SEQ ID NO:6 in control renal cells.

1        40. A method of identifying upregulation of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4  
2    or SEQ ID NO:6 in cells comprising contacting the cells with an antisense probe and measuring  
3    hybridization to RNA within the cell.

1        41. A method of identifying damage or regeneration of renal cells in a subject, comprising  
2    comparing concentration of SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:7 in renal cells, renal  
3    cell fragments or body fluids of the subject to a control level of expression of SEQ ID NO:3,  
4    SEQ ID NO:5 or SEQ ID NO:7 in control renal cells.

1        42. The method of claim 41, wherein the fluid is urine or serum.

1        43. The method of claim 41, wherein the renal cells or renal cell fragments are obtained  
2    from urine sediment of the subject.

1/11

1	GCGGCCGCGTCGACGGTGCCTGTGAGTAAATAGATCAGGGTCTCCTCAC	50
51	AGCACATTCTCCAGGAAGCCGAGCAAACATTAGTGTCTTTACCCAGGA	100
101	GGAAATCTAGGTGTAGAGAGCTCTACGGATCTAAGTTGGATCTGTACC	150
151	CAGTGCTTTTAGGTGTCTTAGACATTCTCAGGAAGATGTAGTCTCT	200
201	GTCACCATGTGTGGCTGAATTCTAGCTCAGTCATCTTATTGTGTTAAG	250
251	GTAGTTGAAGTTAGGAACCAACCAGTATGTCTGTGAGCAGAAGAGTACA	300
301	GTGTCCATCTTGAGGACAAGCTCATCTTACCATAGAGGGCTGGCCTTG	350
351	GCTTAGATTCTACCGAGAACATACTCTTAATGGCTGCCCTCAGTTTCT	400
401	CTGTTGCTGTCTTATTGTGTCTAGGCCAGAAGTCATATGGATGGCTCT	450
451	ATGTGAGCAAGGACCCAGATAGAAGAGTGTATTGGGGAACAGGTTGCC	500
501	CTAACAGAGAGTCCTGTGGATTCATGCAGTCAGGATGAAGACCTGATCA	550
551	GACAGAGTGTGCTGAGTGCCACGGCTAACCAAGAGTGACTTGTCACTGTCC	600
601	TTCAGGTCAACACCATGGTTCAACTCAAGTCTCATTTCAGGCCTCCTG M V Q L Q V F I S G L L	650
651	CTGCTTCTTCCAGGCTCTGTAGATTCTTATGAAGTAGTGAAGGGGGTGGT L L L P G S V D S Y E V V K G V V	700
701	GGGTCACCCCTGTACAATTCCATGTACTTACTCAACACGTGGAGGAATCA G H P V T I P C T Y S T R G G I T	750
751	CAACGACATGTTGGGGCCGGGGCAATGCCATATTCTAGTTGTAAAAT T T C W G R G Q C P Y S S C Q N	800
801	ATACTTATTGGACCAATGGATACCAAGTCACCTATCGGAGCAGCGGTG I L I W T N G Y Q V T Y R S S G R	850
851	ATACAACATAAAGGGCGTATTCAGAAGGAGACGTATCCTTGACAATAG Y N I K G R I S E G D V S L T I E	900
901	AGAACTCTGTTGATAGTGTAGTGGTCTGTATTGTTGCCAGTGGAGATT N S V D S D S G L Y C C R V E I	950
951	CCTGGATGGTTCAACGATCAGAAAATGACCTTTCATTGGAAGTTAAACC P G W F N D Q K M T F S L E V K P	1000
1001	AGAAAATTCCCACAAGTCCTCCAACAAGACCCACAACCTACAAGACCCACAA E I P T S P P T R P T T T R P T T	1050
1051	CCACAAAGGCCACAACATATTCAACAAGATCCACACATGTACCAACATCA T R P T T I S T R S T H V P T S	1100

FIG. 1a

2/11

1101	ACCAAGAGTCTCCACCTCTACTCCAACACCCAGAACAAACACAGACTCACAA T R V S T S T P T P E Q T Q T H K	1150
1151	ACCAAGAAATCACTACATTATGCCCATGAGACAACGTGCTGAGGTGACAG P E I T T F Y A H E T T A E V T E	1200
1201	AAACTCCATCATATACTCCTGCAGACTGGAATGGCACTGTGACATCCTCA T P S Y T P A D W N G T V T S S	1250
1251	GAGGAGGCCTGGAATAATCACACTGTAAGAATCCCTTGAGGAAGCCGCA E E A W N N H T V R I P L R K P Q	1300
1301	GAGAAACCCGACTAAGGGCTCTATGTTGGCATGTCCGTTGCAGCCCTGC R N P T K G F Y V G M S V A A L L	1350
1351	TGCTGCTGCTGCTGCAGCACCGTGGTGTCAACCAGGTACATCATTATA L L L L A S T V V V T R Y I I I	1400
1401	AGAAAAGAAGATGGGCTCTCTGAGCTTGTGCCTTCATGTCTAAAGAG R K K M G S L S F V A F H V S K S	1450
1451	TAGAGCTTGAGAACCGCAGCGATTGTGCATCCCCGAGCTGAAGACAACA R A L Q N A A I V H P R A E D N I	1500
1501	TCTACATTATTGAAGATAGATCTCGAGGTGCAGAACATGAGTCCCAGAGGCC Y I I E D R S R G A E	1550
1551	TTCTGTGGGCCTTCTGCCTGGATTACAGAGATCGTACTGATTCACA	1600
1601	GAGTAAAATACCCATTCCAGCTCCTGGAGATTGTGTTGGTTCTTC	1650
1651	CAGCTGCAGTGGAGAGGGTAACCCCTCACCCGTATATGCAAAACCTCGAG	1700
1701	GTAAACATCATCTTAATTCTGTATCAGCAACACCTCAGTGTCTCCACTC	1750
1751	ACTGCAGCGATTCTCTCAAATGTGAACATTAGAAGTTGTGTTCCCTT	1800
1801	TTGTCCATGTAATCATTGTAATACAAGAATTATTTATCTTGTATTAAAAA	1850
1851	CCATTAATGAGAGGGAAATAGGAATTAAAAGCTGGTGGAAAGGGCCTCCT	1900
1901	GAATTAGAACCTCATGATTGTGTTATCTCTTTATTGTAATTGAA	1950
1951	AATGTTACTTCTATCCTTCCAAGGGCAAAATCATGGGAGCATGGAGGT	2000
2001	TTTAATTGCCCTCATAGATAAGTAGAAGAACAGAGACTAATGCCACCAAT	2050
2051	AGAGGTGGTTATGCTTCTCACAGCTCTGGAAATATGATCATTATTATG	2100
2101	CAGTTGATCTTAGGATGAGGATGGGTTCTTAGGAGGAGAGGTACCATG	2150
2151	GTGAGTGGACCAGGCACACATCAGGGGAAGAAAACAATGGATCAAGGGAT	2200
2201	TGAGTTCATTAGAGCCATTCCACTCCACTTCTGTCTTGATGCTCAGTGT	2250
2251	TCCTAAACTCACCCACTGAGCTCTGAATTAGGTGCAGGGAGGAGACGTGC	2300

FIG. 1b

SUBSTITUTE SHEET (RULE 26)

3/11

2301	AGAAACGAAAGAGGAAAGAAAGGAGAGAGCAGGACACAGGCTTCCTGC	2350
2351	TGAGAGAAGTCCTATTGCAGGTGTGACAGTGTGGACTACCACGGGTT	2400
2401	TCCTTCAGACTTCTAACGTTCTAAATCACTATCATGTGATCATATTATT	2450
2451	TTTAAAATTATTCAGAAAGACACCACATTTCAATAATAAAATCAGTTG	2500
2501	TCACAATTAATAAAATATTTGTTTGCTAAGAAGTAAAAAAAAAAAAAA	2550
2551	AAGTCGACGCCCGC 2566	

**FIG. 1c**

4/11

1	GC GG CC CG CT CG AC GG TG C G C T GT G AGT AA AT AG AT C AG GG CT C C TT C A C	50
51	AG C A C A T T C T C C A G G A A G C C G A G C A A A C A T T A G T G C T A T T T A C C C A G G A	100
101	GG A A A T C T A G G T G T A G A G A G C T C T A C G G A T C T A A G G T C A A C A C C A T G G T T M V	150
151	CA A C T T C A A G T C T T C A T T C A G G C C T C C T G C T G C T T C T T C C A G G C T C T G T Q L Q V F I S G L L L L P G S V	200
201	A G A T T C T T A T G A A G T A G T G A A G G G G T G G T G G G T C A C C C T G T C A C A A T T C D S Y E V V K G V V G H P V T I P	250
251	C A T G T A C T T A C T C A A C A C G T G G A G G A A T C A C A A C G A C A T G T G G G G C C G G C T Y S T R G G I T T T C W G R	300
301	G G G C A A T G C C C A T A T T C T A G T T G T C A A A A T A T A C T T A T T G G A C C A A T G G G Q C P Y S S C Q N I L I W T N G	350
351	A T A C C A A G T C A C C T A T C G G A G C A G C G G T C G A T A C A A C A T A A A G G G G C G T A Y Q V T Y R S S G R Y N I K G R I	400
401	T T T C A G A A G G A G A C G T A T C C T T G A C A A T A G A G A A C T C T G T T G A T A G T G A T S E G D V S L T I E N S V D S D	450
451	A G T G G T C T G T A T T G T T G C C G A G T G G A G A T T C C T G G A T G G T T C A A C G A T C A S G L Y C C R V E I P G W F N D Q	500
501	G A A A A T G A C C T T T C A T T G G A A G T T A A C C A G A A A T T C C C A C A A G T C C T C K M T F S L E V K P E I P T S P P	550
551	C A A C A A G A C C C A C A A C T A C A A G A C C C A C A A C C A A G G C C C A C A A C T A T T T R P T T T R P T T T R P T T I	600
601	T C A A C A A G A T C C A C A C A T G T A C C A A C A T C A A C C A G A G T C T C C A C C T C T A C S T R S T H V P T S T R V S T S T	650
651	T C C A A C A C C A G A A C C A A C A C A G A C T C A C A A A C C A G A A A T C A C T A C A T T T P T P E Q T Q T H K P E I T T F Y	700
701	A T G C C C A T G A G A C A A C T G C T G A G G T G A C A G A A A C T C C A T C A T A T A C T C C T A H E T T A E V T E T P S Y T P	750
751	G C A G A C T G G A A T G G C A C T G T G A C A T C C T C A G A G G G A G G C T G G A A T A A T C A A D W N G T V T S S E E A W N N H	800
801	C A C T G T A A G A A T C C C T T G A G G A A G C C G C A G A G A A A C C C G A C T A A G G G C T T V R I P L R K P Q R N P T K G F	850
851	T C T A T G T G G C A T G T C C G T T G C A G C C C T G C T G C T G C T G C T G C G A G C Y V G M S V A A L L L L L L A S	900
901	A C C G T G G T T G T C A C C A G G T A C A T C A T T A A G A A A G A A G A G A T G G G C T C T C T T V V V T R Y I I I R K K M G S L	950

**FIG. 2a**  
**SUBSTITUTE SHEET (RULE 26)**

5/11

951	GAGCTTGTGCCTCCATGTCTCTAAGAGTAGAGCTTCAGAACGCAG S F V A F H V S K S R A L Q N A A	1000
1001	CGATTGTGCATCCCCGAGCTGAAGACAACATCTACATTATTGAAGATAGA I V H P R A E D N I Y I I E D R	1050
1051	TCTCGAGGTGCAGAAATGAGTCCCAGAGGCCCTCTGTGGGGCCTCTGCCT S R G A E	1100
1101	GGGATTACAGAGATCGTGAUTTACACAGAGTAAAATACCCATTCCAG	1150
1151	CTCCTGGGAGATTTGTGTTGGTTCTCCAGCTGCAGTGGAGAGGGTA	1200
1201	ACCCCTCTACCCCTGTATATGCAAAACTCGAGGTTAACATCATCCTAATTCT	1250
1251	TGTATCAGCAACACCTCAGTGTCTCCACTCACTGCAGCGATTCTCTCAAA	1300
1301	TGTGAACATTTAGAAGTTGTGTTCTTTGTCCATGTAATCATTGGT	1350
1351	AATAACAAGAATTTATCTTGTATTAAAACCATTAAATGAGAGGGAAATA	1400
1401	GGAATTAAAAGCTGGTGGGAAGGGCCTCCTGAATTAGAAGCACTTCATG	1450
1451	ATTGTGTTATCTCTTTATTGTAATTGAAATGTTACTCTATCCTTC	1500
1501	CAAGGGCAAATCATGGGAGCATGGAGGTTTAATTGCCCTCATAGATA	1550
1551	AGTAGAAGAAGAGAGTCTAATGCCACCAATAGAGGTGGTATGCTTCTC	1600
1601	ACAGCTCTGGAAATATGATCATTATTATGCAGTTGATCTTAGGATGAGG	1650
1651	ATGGGTTCTTAGGAGGAGGGTTACCATGGTGAGTGGACCAGGCACACA	1700
1701	TCAGGGGAAGAAAACAATGGATCAAGGGATTGAGTCATTAGGCCATT	1750
1751	CCACTCCACTCTGTCTTGATGCTCAGTGTCTAAACTCACCCACTGAG	1800
1801	CTCTGAATTAGGTGCAGGGAGGAGACGTGCAGAAACGAAAGAGGAAAGAA	1850
1851	AGGAGAGAGAGCAGGACACAGGCTTCTGCTGAGAGAAGTCCTATTGCAG	1900
1901	GTGTGACAGTGTGTTGGACTACCACGGTTCTCAGACTTCTAAGTT	1950
1951	CTAAATCACTATCATGTGATCATATTATTTAAAATTATTCAGAAAG	2000
2001	ACACCACATTTCAATAATAATCAGTTGTACAATTAAATAAAATATT	2050
2051	TGTTTGCTAAGAAGTAAAAGTCGACGCCGC 2084	

FIG. 2b

6/11

1	GCGGCCGCGTCGACTCGCAGGAGGCCGGCACTCTGACTCCTGGTGGATGG	50
51	GACTAGGGAGTCAGAGTCAGCCCTGACTGGCTGAGGGCGGGCGCTCCGA	100
101	GTCAGCATGGAAAGTCTCTCGGGGTCCTGGTATTTCTGCTGCTGGCTGC M E S L C G V L V F L L A A	150
151	AGGACTGCCGCTCCAGGCGCCAAGCGGTTCCGTGATGTGCTGGGCCATG G L P L Q A A K R F R D V L G H E	200
201	AGCAGTATCCGGATCACATGAGGGAGAACAAACCAATTACGTGGCTGGTCT Q Y P D H M R E N N Q L R G W S	250
251	TCAGATGAAAATGAATGGGATGAACAGCTGTATCCAGTGTGGAGGAGGG S D E N E W D E Q L Y P V W R R G	300
301	AGAGGGCAGATGGAAGGACTCCTGGGAAGGAGGCCGTGTGCAGGCAGCCC E G R W K D S W E G G R V Q A A L	350
351	TAACCAGTGATTCACCGGCCTTGGTGGTTCCAATATCACCTCGTAGTG T S D S P A L V G S N I T F V V	400
401	AACCTGGTGTCCCCAGATGCCAGAACAGGAAGATGCCAACGGCAATATCGT N L V F P R C Q K E D A N G N I V	450
451	CTATGAGAGGAACTGCAGAAGTGATTGGAGCTGGCTTCTGACCCGTATG Y E R N C R S D L E L A S D P Y V	500
501	TCTACAACGGACACAGGGCAGACGATGAGGACTGGGAAGACAGCACC Y N W T T G A D D E D W E D S T	550
551	AGCCAAGGCCAGCACCTCAGGTTCCCCGACGGGAAGGCCCTCCCTCGCCC S Q G Q H L R F P D G K P F P R P	600
601	CCACGGACGGAAGAAATGGAACCTCGTCTACGTCTTCCACACACTGGTC H G R K K W N F V Y V F H T L G Q	650
651	AGTATTTCAAAAGCTGGTCGGTGTTCAGCACGAGTTCTATAAACACA Y F Q K L G R C S A R V S I N T	700
701	GTCAACTGACAGTTGCCCTCAGGTATGGAAGTGATTGTCTTCGAAG V N L T V G P Q V M E V I V F R R	750
751	ACACGGCCGGCATACTTCCATCTCAAAGTGAAAGACGTGTATGTGA H G R A Y I P I S K V K D V Y V I	800
801	TAACAGATCAGATCCCTATATTCGTGACCATGTACCAAGAAGAACCGG T D Q I P I F V T M Y Q K N D R	850
851	AACTCGTCTGATGAAACCTTCCTCAGAGACCTCCCCATTTCCTCGATGT N S S D E T F L R D L P I F F D V	900
901	CCTCATTACGATCCCAGTCATTCTCAACTACTCTGCCATTCTACAC L I H D P S H F L N Y S A I S Y K	950

FIG. 3a

7/11

951	AGTGGAACTTGGGGACAACACTGGCCTGTTGTCTCCAACAATCACACT W N F G D N T G L F V S N N H T	1000
1001	TTGAATCACACGTATGTGCTCAATGGAACCTCAACTTAACCTCACCGT L N H T Y V L N G T F N F N L T V	1050
1051	GCAAACACTGCAGTGCCGGGACCATGCCCTCACCCACACCTCGCCTCTT Q T A V P G P C P S P T P S P S S	1100
1101	CTTCGACTTCTCCTCGCCTGCATCTCGCCTCACCCACATTATCAACA S T S P S P A S S P S P T L S T	1150
1151	CCTAGTCCCTTTAATGCCACTGGCCACAAATCCATGGAGCTGAGTGA P S P S L M P T G H K S M E L S D	1200
1201	CATTTCCAATGAAAATGCCGAATAAACAGATATGGTTACTTCAGAGCCA I S N E N C R I N R Y G Y F R A T	1250
1251	CCATCACAAATTGTAGATGGAATCCTAGAAGTCAACATCATCCAGGTAGCA I T I V D G I L E V N I I Q V A	1300
1301	GATGTCCAATCCCCACACCGCAGCCTGACAACACTCACTGATGGACTTCAT D V P I P T P Q P D N S L M D F I	1350
1351	TGTGACCTGCAAAGGGGCCACTCCCACGGAAGCCTGTACGATCATCTCTG V T C K G A T P T E A C T I I S D	1400
1401	ACCCCACCTGCCAGATGCCAGAACAGGGTGTGCAGCCGGTGGCTGTG P T C Q I A Q N R V C S P V A V	1450
1451	GATGAGCTGTGCCTCCTGTCCGTGAGGAGAGCCTCAATGGTCCGGCAC D E L C L L S V R R A F N G S G T	1500
1501	GTACTGTGTGAATTCTACTCTGGAGACGATGCAAGCCTGGCCCTCACCA Y C V N F T L G D D A S L A L T S	1550
1551	GCGCCCTGATCTCTATCCCTGGCAAAGACCTAGGCTCCCTCTGAGAAC A L I S I P G K D L G S P L R T	1600
1601	GTGAATGGTGTCTGATCTCATTGGCTGCCATGTTGTCAACC V N G V L I S I G C L A M F V T M	1650
1651	GGTACCATCTGCTGTACAAAAACACAAGACGTACAAGCCAATAGGAA V T I L L Y K K H K T Y K P I G N	1700
1701	ACTGCACCAGGAACGTGGTCAAGGGCAAAGGCCTGAGTGTGTTCTCAGC C T R N V V K G K G L S V F L S	1750
1751	CATGCAAAGCCCCGTTCTCCGAGGAGACGGGAGAAGGATCCACTGCT H A K A P F S R G D R E K D P L L	1800
1801	CCAGGACAAGCCATGGATGCTCTAAGTCTCACTCTCACTTCTGACTGGG Q D K P W M L	1850
1851	AACCCACTCTGTGCATGTATGTGAGCTGTGCAGAAGTACATGACTGG 1900	

## FIG. 3b

SUBSTITUTE SHEET (RULE 26)

8/11

1901	TAGCTGTTGTTTCTACGGATTATTGTAAAATGTATATCATGGTTAGGG	1950
1951	AGCGTAGTTAATTGGCATTAGTGAAGGGATGGGAAGACAGTATTCTT	2000
2001	CACATCTGTATTGTGGTTTTATACTGTTAATAGGGTGGGCACATTGTGT	2050
2051	CTGAAGGGGGAGGGGGAGGTCACTGCTACTTAAGGTCCCTAGGTTAACTGG	2100
2101	GAGAGGATGCCCAAGGCTCCTAGATTCTACACAAGATGTGCCTGAACC	2150
2151	CAGCTAGTCCTGACCTAAAGGCCATGTTCATCAACTCTATCTCAGCTCA	2200
2201	TTGAACATACCTGAGCACCTGATGGAATTATAATGGAACCAAGCTTGTG	2250
2251	TATGGTGTGTGTACATAAGATACTCATTAAAAAGACAGTCTATTAA	2300
2301	AAA 2303	

## FIG. 3c

9/11

1	ATGCATCCTCAAGTGGTCATCTTAAGCCTCATCCTACATCTGGCAGATT M H P Q V V I L S L I L H L A D S	50
51	TGTAGCTGGTTCTGTAAAGGTTGGAGAGGCAGGTCCATCTGTACAC V A G S V K V G G E A G P S V T L	100
101	TACCCCTGCCACTACAGTGGAGCTGTACATCAATGTGCTGGAATAGAGGC P C H Y S G A V T S M C W N R G	150
151	TCATGTTCTCTATTCACATGCCAAAATGGCATTGTCTGGACCAATGGAAC S C S L F T C Q N G I V W T N G T	200
201	CCACGTCACCTATCGGAAGGACACACGCTATAAGCTATTGGGGACCTTT H V T Y R K D T R Y K L L G D L S	250
251	CAAGAAGGGATGTCTCTTGACCATAAGAAAATACAGCTGTGTCTGACAGT R R D V S L T I E N T A V S D S	300
301	GGCGTATATTGTTGCCGTGTTGAGCACCGTGGGTGGTTCAATGACATGAA G V Y C C R V E H R G W F N D M K	350
351	AATCACCGTATCATTGGAGATTGTGCCACCCAAGGTACGACTACTCAA I T V S L E I V P P K V T T T P I	400
401	TTGTCACAACGTCCAACCGTCACGACTGTTGAACGAGCACCACTGTT V T T V P T V T T V R T S T T V	450
451	CCAACGACAACGACTGTTCCAACGACAACGACAACATGAGCAT P T T T V P T T T V P T T M S I	500
501	TCCAACGACAACGACTGTTCCGACGACAATGACTGTTCAACGACAACGA P T T T V P T T M T V S T T T S	550
551	GCGTTCCAACGACAACGAGGATTCCAACAAACAACAAGTGTCCAGTGACA V P T T T S I P T T T S V P V T	600
601	ACAACGGTCTCTACCTTGTCTCCAATGCCTTGCCCCAGGCAGAACCA T T V S T F V P P M P L P R Q N H	650
651	TGAACCACTAGCCACTTCACCATCTTCACCTCAGCCAGCAGAAACCCACC E P V A T S P S S P Q P A E T H P	700
701	CTACGACACTGCAGGGAGCAATAAGGAGAGAACCCACCACTCACCATTG T T L Q G A I R R E P T S S P L	750
751	TACTCTTACACAACAGATGGAATGACACCGTGACAGAGTCAGATGG Y S Y T T D G N D T V T E S S D G	800
801	CCTTTGGAATAACAATCAAACACTCAAACGTCTAGAACATAGTCTACTGA L W N N N Q T Q L F L E H S L L T	850
851	CGGCCAATACCACTAAAGGAATCTATGCTGGAGTCTGTATTCGTCTTG A N T T K G I Y A G V C I S V L	900

FIG. 4a

10/11

901	GTGCTTCTTGCTCTTTGGGTGTCATCATTGCCAAAAGTATTCTTCAA V L L A L L G V I I A K K Y F F K	950
951	AAAGGAGGTTCAACAACTAAGACCCCATAAATCCTGTATACATCAAAGAG K E V Q Q L R P H K S C I H Q R E	1000
1001	AA 1002	

**FIG. 4b**

11/11

**FIG. 5**

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 97/09303

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>					
IPC 6 C12N15/12 C07K14/47 C12N15/62 C07K16/18 A61K38/16 G01N33/50 C12Q1/68 C12N1/21 C12N5/10 C12N5/12 A61K48/00					
According to International Patent Classification (IPC) or to both national classification and IPC					
<b>B. FIELDS SEARCHED</b>					
Minimum documentation searched (classification system followed by classification symbols)					
IPC 6 C12N C07K A61K G01N C12Q					
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched					
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)					
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>					
Category *	Citation of document, with indication, where appropriate, of the relevant passages				Relevant to claim No.
X	WO 96 04376 A (US HEALTH) 15 February 1996 see page 39 - page 41 see page 43 ----- XETERMAN A.M.J. ET AL: "nmb, a novel gene, is expressed in low-metastatic human melanoma cell lines and xenografts" INT. J. CANCER, vol. 60, 1995. pages 73-81, XP002044250 see the whole document -----				1  1
<input type="checkbox"/> Further documents are listed in the continuation of box C.			<input checked="" type="checkbox"/> Patent family members are listed in annex.		
<b>* Special categories of cited documents :</b> *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed					
*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *S* document member of the same patent family					
4 Date of the actual completion of the international search			Date of mailing of the international search report		
22 October 1997			04.11.97		
Name and mailing address of the ISA			Authorized officer		
European Patent Office, P.O. Box 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax. 31 651 epo nl Fax: (+31-70) 340-3016			Espen, J		

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 97/09303

### Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
**see FURTHER INFORMATION sheet PCT/ISA/210**
2.  Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

#### Remark on Protest

- The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 97 09303

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Remark : Although claims 22-36 are directed to a method of treatment of the human/animal body , and although claims 38-39, and in part 37,40,41 are directed to a diagnostic method practised on the human/animal body , the search has been carried out and based on the alleged effects of the compound/composition.

**INTERNATIONAL SEARCH REPORT**International Application No  
**PCT/US 97/09303**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9604376 A	15-02-96	US 5622861 A AU 3238995 A	22-04-97 04-03-96